B. G. TEUBNERS LEHRBÜCHER
DER MATHEMATISCHEN WISSENSCHAFTEN IX 1

E. CZUBER

WAHRSCHEINLICHKEITSRECHNUNG

I

B. G. Teubners Sammlung von Lehrbüchern auf dem Gebiete der Mathematischen Wissenschaften mit Einschluß ihrer Anwendungen.

gr. 8.  geb.

Diese Sammlung bietet in einzelnen in sich abgeschlossenen Werken zusammenfassende Darstellungen der wichtigsten Abschnitte der mathematischen Wissenschaften und deren Anwendungen. Im einzelnen wollen diese Werke in ihrer ausführlichen, neben der rein wissenschaftlichen auch pädagogische Momente berücksichtigenden Darstellung die Möglichkeit zu selbständigem und umfangreichen Quellenstudien unabhängigem Eindringen in die verschiedenen Disziplinen geben; in ihrer Gesammtheit aber sollen sie durch ihre eingehenden literarischen und historischen Nachweise ein genaues Bild von der modernen Entwicklung und von dem gegenwärtigen Stande der mathematischen Wissenschaften und ihrer Anwendungen darbieten.

Bisher erschienen folgende Bände:

P. Bachmann, niedere Zahlentheorie. 2 Bände. I. Band. X, 402 S. 1902. n. ℳ 14.—. [Bd. X, 1.]

E. Blaschke, Vorlesungen über mathematische Statistik. Die Lehre von den statistischen Maßzahlen. Mit 17 Textfiguren und 5 Tafeln. VIII, 268 S. 1906. n. ℳ 7.40. [Bd. XXIII.]

H. Bruns, Wahrscheinlichkeitsrechnung und Kollektivmaßlehre. VIII, 310 S. und Anhang 18 S. 1906. n. ℳ 8.40. [Bd. XVII.]

G. H. Bryan, Thermodynamics. An introductory Treatise dealing mainly with first Principles and their direct Applications. Mit 26 Fig. XIV, 204 S. 1907. n. ℳ 7.—. (Englisch.) [Bd. XXI.]

E. Czuber, Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung. 2. Auflage in 2 Bänden. I. Band: Wahrscheinlichkeitstheorie. Fehlerausgleichung. Kollektivmaßlehre. Mit 18 Fig. im Text. X, 410 S. 1908. n. ℳ 12.—. [Bd. IX, 1.]

L. E. Dickson, linear Groups with an Exposition of the Galois Field theory. X, 312 S. 1901. n. ℳ 12.—. (Englisch.) [Bd. VI.]

O. Fischer, theoretische Grundlagen für eine Mechanik der lebenden Körper mit speziellen Anwendungen auf den Menschen, sowie auf einige Bewegungsvorgänge an Maschinen. In möglichst elementarer und anschaulicher Weise dargestellt. Mit 67 Fig. u. 4 Taf. X, 372 S. 1906. n. ℳ 14.—. [Bd. XXII.]

A. Gleichen, Lehrbuch der geometrischen Optik. Mit 251 Fig. XIV, 511 S. 1902. n. ℳ 20.—. [Bd. VIII.]

A. Krazer, Lehrbuch der Thetafunktionen. Mit 10 Figuren. XXIV, 509 S. 1903. n. ℳ 24.—. [Bd. XII.]

H. Lamb, Lehrbuch der Hydrodynamik. Deutsch von Joh. Friesendorff. Mit 79 Fig. XVI, 787 S. 1907. n. ℳ 20.—. [Bd. XXVI.]

R. von Lilienthal, Vorlesungen über Differentialgeometrie. In 2 Bänden. I. Band: Kurventheorie. Mit 26 Fig. VI, 368 S. n. ℳ 12.—. [Bd. XXVIII, 1.]

TS 30/108

G. Loria, spezielle algebraische und transzendente ebene Kurven. Theorie und Geschichte. Deutsch von Fr. Schütte. Mit 174 Fig. auf 17 lithogr. Tafeln. XXI, 744 S. 1902. n. M 28.—. [Bd. V.]

—— Vorlesungen über darstellende Geometrie. Deutsch von Fr. Schütte. In 2 Teilen. I. Teil: Die Darstellungsmethoden. Mit 163 Figuren. XI, 219 S. 1906. n. M 6.80. [Bd. XXV, 1.]

A. E. H. Love, Lehrbuch der Elastizität. Deutsch unter Mitwirkung des Verfassers von A. Timpe. Mit 75 Abbildungen. XVI, 664 S. 1907. n. M 16.— [Bd. XXIV.]

E. Netto, Lehrbuch der Kombinatorik. VIII, 260 S. 1901. n. M 9.—. [Bd. VII.]

W. F. Osgood, Lehrbuch der Funktionentheorie. 2 Bände. I. Band. Mit 150 Figuren. XII, 642 S. 1907. n. M 15.60. [Bd. XX, 1.]

E. Pascal, die Determinanten. Eine Darstellung ihrer Theorie und Anwendungen mit Rücksicht auf die neueren Forschungen. Berechtigte deutsche Ausgabe von H. Leitzmann. XVI, 266 S. 1900. M 10.—. [Bd. III.]

Fr. Pockels, Lehrbuch der Kristalloptik. Mit 168 Figuren und 6 Doppeltafeln. X, 519 S. 1906. n. M 16.—. [Bd. XIX.]

D. Seliwanoff, Lehrbuch der Differenzenrechnung. VI, 92 S. 1904. n. M 4.— [Bd. XIII.]

O. Staude, analytische Geometrie des Punktes, der geraden Linie und der Ebene. Ein Handbuch zu den Vorlesungen und Übungen über analytische Geometrie. Mit 387 Figuren. VIII, 447 S. 1905. n. M 14.—. [Bd. XVI.]

O. Stolz und J. A. Gmeiner, theoretische Arithmetik. 2., umgearbeitete Aufl. ausgewählter Abschnitte der „Vorlesungen über allgemeine Arithmetik" von O. Stolz. XI, 402 S. 1902. n. M 10.60. [Bd. IV.]

—— Einleitung in die Funktionentheorie. 2., umgearbeitete und vermehrte Auflage der von den Verfassern in der „Theoretischen Arithmetik" nicht berücksichtigten Abschnitte der „Vorlesungen über allgemeine Arithmetik" von O. Stolz. Mit 21 Fig. X, 598 S. 1905. n. M 15.—. [Bd. XIV.]

R. Sturm, die Lehre von den geometrischen Verwandtschaften. 4 Bände. I. Band: Die Verwandtschaften zwischen Gebilden erster Stufe. XII, 415 S. 1908. n. M 18.—. [Bd. XXVII, 1.]

H. E. Timerding, Geometrie der Kräfte. [illegible] 1908. n. M 16.—. [Bd. I.]

J. G. Wallentin, Einleitung in die theoretische Elektrizitätslehre. Mit [illegible] Fig. X, 444 S. 1904. n. M 12.—. [Bd. XV.]

E. von Weber, Vorlesungen über das Pfaffsche Problem und die Theorie der partiellen Differentialgleichungen erster Ordnung. XI, 622 S. 1900. n. M 24.—. [Bd. II.]

A. G. Webster, the Dynamics of particles and of rigid, elastic, and fluid Bodies, being Lectures on mathematical Physics. Mit 173 Fig. XII, [illegible] S. 1904. n. M 14.—. (English.) [Bd. XI.]

E. J. Wilczynski, projective differential geometry of Curves and ruled Surfaces. VIII, 298 S. 1906. n. M 10.—. (English.) [Bd. XVIII.]

Unter der Presse:

E. Czuber, Wahrscheinlichkeitsrechnung. 2. Aufl. 2 Bände. [illegible]

H. A. Lorentz, on the Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat. [illegible]

G. Loria, Vorlesungen über darstellende Geometrie. [illegible]

R. Sturm, die Lehre von den geometrischen Verwandtschaften. 4 Bände. [illegible]

B. G. TEUBNERS SAMMLUNG VON LEHRBÜCHERN
AUF DEM GEBIETE DER
MATHEMATISCHEN WISSENSCHAFTEN
MIT EINSCHLUSS IHRER ANWENDUNGEN
BAND IX,1

WAHRSCHEINLICHKEITSRECHNUNG

UND IHRE ANWENDUNG AUF FEHLERAUSGLEICHUNG STATISTIK UND LEBENSVERSICHERUNG

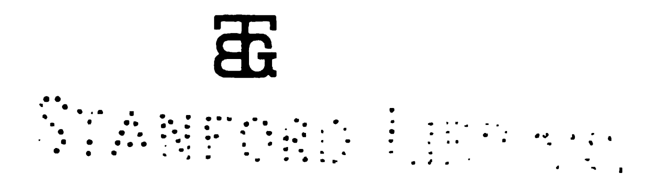
VON

EMANUEL CZUBER
O. Ö. PROFESSOR AN DER TECHNISCHEN HOCHSCHULE IN WIEN

ERSTER BAND
WAHRSCHEINLICHKEITSTHEORIE · FEHLERAUSGLEICHUNG
KOLLEKTIVMASSLEHRE

MIT 18 FIGUREN IM TEXT

ZWEITE, SORGFÄLTIG DURCHGESEHENE
UND ERWEITERTE AUFLAGE



LEIPZIG UND BERLIN
DRUCK UND VERLAG VON B. G. TEUBNER
1908

Vorwort.

Die Wahrscheinlichkeitsrechnung hat längst aufgehört, eine bloße Theorie der Glücksspiele zu sein. Man hat einerseits den hohen Bildungswert ihrer Begriffskonstruktionen und Schlußweisen erkannt und andrerseits den Kreis jener konkreten Materien, auf welche sich ihre Methoden mit Berechtigung anwenden lassen, immer bestimmter gezogen; so ist die Wahrscheinlichkeitsrechnung zu einem wohlbegründeten und praktisch wichtigen Zweige der angewandten Mathematik ausgestaltet worden.

Sie nach den beiden erwähnten Richtungen innerhalb vorgezeichneter Schranken zur Darstellung zu bringen, ist das Ziel, das ich mir bei Abfassung dieses Buches gesteckt habe.

Der grundlegende erste Teil geht auf die fundamentalen Fragen soweit ein, als es zur Vorbereitung jener Kritik erforderlich ist, die geübt werden soll, wenn man die Wahrscheinlichkeitsrechnung auf ein Gebiet der Wirklichkeit anzuwenden sich anschickt. Eine reiche Auswahl von Problemen, darunter die klassischen, soll mit der eigenartigen Denkweise, mit der richtigen Handhabung der Sätze und Rechnungsmethoden vertraut machen.

Für alle Gebiete der Forschung, wo die Resultate aus messenden Beobachtungen abgeleitet werden, ist die auf die Wahrscheinlichkeitstheorie sich stützende Ausgleichungsrechnung zu einem unentbehrlichen Instrument geworden, dazu bestimmt, die Gewinnung der Resultate methodisch zu leiten und für ihre Verläßlichkeit einen Maßstab zu liefern. Diese Seite der Anwendung behandelt der zweite Teil, der in eine kurzgefaßte Theorie der Beobachtungsfehler und in die Erledigung der Hauptprobleme der Kombination von Beobachtungen zerfällt. Erläuternde Beispiele sind in zureichender Zahl eingefügt.

Ein verhältnismäßig junger Zweig der Anwendung ist die mathematische Statistik, die den Gegenstand des dritten Teiles bildet. Neben den Methoden zur Kritik statistischer Resultate werden die Probleme der Sterblichkeits- und Invaliditätsmessung zur Sprache gebracht. Die Hilfsmittel der formalen Bevölkerungstheorie kommen dabei zu weitgehender Anwendung. Die reichliche Vorführung von Erfahrungsmaterial wird dazu beitragen, die Ausführungen dieses Teiles zu beleben.

Im vierten Teile wird die Lebensversicherungsrechnung auf wahrscheinlichkeitstheoretischer Grundlage zum Vortrag gebracht. Im Gegensatze zu dem älteren Verfahren, das jede Versicherungsart für sich betrachtet und

immer wieder von Grund auf in Angriff nimmt, ist hier im Interesse der zusammenfassenden Behandlung eine eingehende Darstellung der mannigfachen Versicherungswerte an die Spitze gestellt; denn alle andern Fragen, wie Prämienbestimmung, Rückgewähr der Prämien, Reserveberechnung, Risiko usw. führen auf das Rechnen mit Versicherungswerten zurück. Angewendet ist das Bezeichnungssystem des Text Book, dieses Hauptwerkes der Lebensversicherungsrechnung, ein System, das am ehesten Aussicht hat, allgemein adoptiert zu werden. Tabellen sind in solchem Umfange aufgenommen, als es notwendig erscheint, die Auswertung der Formeln zu erläutern und mit speziellen Zahlwerten dieses Gebietes einigermaßen bekannt zu machen.

Betreffs weitergehender literarischer und historischer Nachweisungen darf ich auf meine Schrift „Die Entwicklung der Wahrscheinlichkeitstheorie und ihrer Anwendungen“ im VII. Bande der Jahresberichte der Deutschen Mathematiker-Vereinigung hinweisen.

Wien, 9. November 1902.

E. Czuber.

Begleitwort zur zweiten Auflage.

Bei der Bearbeitung der zweiten Auflage dieses Teils habe ich mancherlei Neuerungen im einzelnen getroffen, die mir förderlich schienen, so die Darstellung der Wahrscheinlichkeitssätze in Form von Funktionalgleichungen, die Heranziehung des Begriffs der relativen Wahrscheinlichkeit, der Mengenlehre. Des weiteren war ich darauf bedacht, die Grundfragen, welche die philosophische Seite des Gegenstandes betreffen, tiefer zu fassen. Ein Kapitel über die Kollektivmaßlehre, die, von G. Th. Fechner begründet, durch die neueren Arbeiten von G. F. Lipps und H. Bruns wesentlich gefördert wurde, durfte nicht mehr fehlen; ich habe die theoretischen Grundlagen dieses jüngsten Zweiges so knapp als möglich dargestellt, hingegen auf die praktische Anwendung durch Vorführung mehrerer, darunter auch größerer Beispiele vorzubereiten gesucht.

Für die von verschiedenen Seiten eingelangten Berichtigungen von Druckversehen und Nachrechnungen spezieller Zahlenwerte, die ich bei der Revision benutzen konnte, sage ich hier meinen Dank.

Ebenso danke ich Herrn Prof. Bruns für die freundliche Erlaubnis, seine Tafeln zur Kollektivmaßlehre aufnehmen zu dürfen.

Wien, 23. Februar 1908.

Der Verfasser.

Inhalt.

Erster Teil.

Wahrscheinlichkeitstheorie.

I. Abschnitt. Grundlagen der Wahrscheinlichkeitsrechnung.

§ 1. Entwicklung der Grundbegriffe.

§ 2. Direkte Wahrscheinlichkeitsbestimmung.

§ 3. Indirekte Wahrscheinlichkeitsbestimmung.

§ 4. Geometrische Wahrscheinlichkeit.

II. Abschnitt. Wahrscheinlichkeiten, betreffend die Ergebnisse wiederholter Beobachtungen.

§ 1. Das Theorem von Bernoulli.

§ 2. Das Theorem von Poisson.

§ 3. Das Gesetz der großen Zahlen.

III. Abschnitt. Wahrscheinlichkeiten auf Grund der Erfahrung.

§ 1. Wahrscheinlichkeit der möglichen Ursachen eines beobachteten Ereignisses.

§ 2. Wahrscheinlichkeit künftiger Ereignisse auf Grund von Beobachtungen.

IV. Abschnitt. Bewertung von Vor- und Nachteilen, welche an zufällige Ereignisse geknüpft sind.

§ 1. Die mathematische Erwartung.

§ 2. Das mathematische Risiko.

§ 3. Die moralische Erwartung.

Zweiter Teil.

Ausgleichungsrechnung.

I. Abschnitt. Theorie der Beobachtungsfehler.

§ 1. Das Fehlergesetz.

§ 3. Bedingte Beobachtungen.

Dritter Teil.

Kollektivmaßlehre.

§ 1. Kollektivgegenstände und die Mittel zu ihrer Beschreibung.

§ 2. Methodische Durchführung der Rechnungen.

§ 3. Anwendungen.

Tafeln.

Erster Teil.

Wahrscheinlichkeitstheorie.

I. Abschnitt. Grundlagen der Wahrscheinlichkeitsrechnung.

§ 1. Entwicklung der Grundbegriffe.

1. Das hypothetische Urteil. Unser Denken vollzieht sich in der Bildung von Begriffen und in ihrer Verknüpfung zu Urteilen.

Unter den verschiedenen Urteilsformen nimmt hier vor allem das hypothetische Urteil unsere Aufmerksamkeit in Anspruch. Sein Wesen besteht darin, daß es an einen Vordersatz einen Nachsatz als notwendige Folge knüpft; seine typische Form lautet: „Wenn A gilt, so gilt B“, oder „Wenn A geschieht, so tritt B ein“. Es sagt also aus, daß mit der Verwirklichung des Vordersatzes stets unabweislich das verbunden ist, was im Nachsatze behauptet wird. Vordersatz und Nachsatz stehen also im Verhältnis von *Grund* und *Folge*; das Prädikat des Urteils ist *Notwendigkeit* (Gewißheit).

Der so geartete Zusammenhang zwischen den beiden Teilen des Urteils wird entweder aus gewissen grundlegenden Urteilen (Axiomen) durch logische Schlüsse erwiesen oder aus einer ausnahmlos gemachten Beobachtung gefolgert.

Die Mathematik drückt ihre Wahrheiten in hypothetischen Urteilen, den Lehrsätzen, aus und gibt für sie Beweise. Das Streben der Naturwissenschaften geht dahin, die Vorgänge der Außenwelt immer mehr in unverbrüchliche Regeln zu fassen, die dann nach Art der mathematischen Lehrsätze auch in der Form hypothetischer Urteile ausgesprochen werden können.

„Wenn eine dekadische Zahl, die an der letzten Stelle 0 oder 5 hat, durch 5 dividiert wird, so geht die Division ohne Rest auf“; — „Wenn man durch eine Ecke eines Dreiecks eine Gerade derart legt, daß sie den an dieser Ecke liegenden Winkel durchsetzt, so schneidet die Gerade die der Ecke gegenüberliegende Seite des Dreiecks“ — sind hypothetische Urteile, deren Wahrheit die Arithmetik, beziehungsweise die Geometrie aus ihren Grundlagen durch logische Schlüsse begründet. Dagegen beruht das Urteil: „Wenn ein Gefäß eine Flüssig-

keit enthält, und ich bringe in dem von ihr berührten Teil der Seitenwand eine Öffnung an, so tritt durch diese die Flüssigkeit aus" — auf einer ausnahmslos gemachten Wahrnehmung.

2. Das hypothetisch-disjunktive Urteil. Wenn aus der Verwirklichung einer Voraussetzung eine von mehreren einander ausschließenden Folgen notwendig hervorgehen muß, so heißt das Urteil, welches diesen Sachverhalt behauptet, ein hypothetisch-disjunktives. Die Unterscheidung und Aufzählung der Folgen bildet die Disjunktion, die einzelnen Folgen heißen ihre Glieder. Die typische Form des disjunktiven Urteils kann in die Worte gefaßt werden: „Wenn A geschieht, so tritt entweder B oder C oder D ... ein." Zwischen der Voraussetzung und der Gesamtheit der Glieder der Disjunktion besteht wie bei dem hypothetischen Urteil das Verhältnis der Notwendigkeit; dagegen steht das einzelne Glied der Disjunktion oder eine Gruppe von Gliedern zu der Voraussetzung im Verhältnisse der *Möglichkeit*, die das eigentliche Prädikat des disjunktiven Urteils ausmacht: Die Verwirklichung der Voraussetzung *kann* das bezeichnete Glied oder eines aus der gewählten Gruppe zur Folge haben, es kann sich aber auch ein anderes Glied einstellen.

„Wenn eine dekadische Zahl durch 5 dividiert wird, so geht die Division auf oder es bleibt ein Rest übrig"; — „Wenn man durch eine Ecke eines Dreiecks eine Gerade legt, so schneidet sie die gegenüberliegende Seite oder sie geht an ihr vorüber"; — „Wenn man in der Seitenwand eines Flüssigkeit enthaltenden Gefäßes eine Öffnung anbringt, so tritt durch diese Flüssigkeit aus oder es geschieht dies nicht" — sind Beispiele disjunktiver Urteile mit zweigliedriger Disjunktion. In dem ersten kann die Disjunktion weiter geführt werden durch Auseinanderhaltung der Reste, welche auftreten können.

Die vorgeführten Beispiele zeigen eine bestimmte Struktur, die sich folgendermaßen beschreiben läßt. Man kann von einem *Bereich* der Voraussetzungen oder Bedingungen sprechen und diese in Teile zerlegen derart, daß an den einen die eine der Folgen, an den andern die andere Folge mit Notwendigkeit gebunden ist; mit andern Worten: man kann die obigen hypothetisch-disjunktiven Urteile in hypothetische auflösen. In dem ersten Beispiel ist der Bereich der allgemeinen Bedingungen durch die Gesamtheit der dekadischen Zahlen vertreten und scheidet sich in solche, die 0 oder 5, und in solche, die eine andere Ziffer an der niedrigsten Stelle haben: mit der ersten Teilgesamtheit ist das Aufgehen der Division, mit der zweiten ein Rest notwendig verbunden. In dem zweiten Beispiel stellt die Gesamtheit der durch die betreffende Ecke gehenden Geraden den Bereich der Bedingungen dar und zerfällt in jene Geraden, die den Winkel an der Ecke durchsetzen, und in solche, die außerhalb desselben bleiben: die ersteren schneiden notwendig, die letzteren gehen notwendig vorbei. In dem

dritten Beispiel vertritt die ganze Seitenwand den Bereich, innerhalb dessen die allgemeine Voraussetzung realisiert werden kann, und scheidet sich in einen von der Flüssigkeit berührten und einen freien Teil: erfolgt die Realisierung in dem ersten, so ist der Austritt der Flüssigkeit die notwendige Folge; dem andern Teil entspricht ebenso das Ausbleiben dieser Erscheinung.

In allen drei Fällen ist man auch in der Lage, das *quantitative Verhältnis* der Umfänge der Teilbereiche zum Bereich der allgemeinen Bedingungen anzugeben. So ist das Verhältnis der Gesamtheit der mit 0 oder 5 schließenden Zahlen zur Gesamtheit aller Zahlen mit 1 : 5 zu bemessen, weil von je zehn Zahlen, die sich nur in der Endziffer unterscheiden, zwei zur ersteren Gesamtheit gehören. Hat ferner in einem konkreten Dreieck der Winkel jener Ecke, durch welche die Gerade gelegt werden soll, die Größe α, so ist $\alpha : \pi$ das Verhältnis der Gesamtheit der die Gegenseite schneidenden Geraden zur Gesamtheit aller Geraden, welche durch die Ecke gelegt werden können. In dem Beispiel mit dem Gefäße bezeichnet das Verhältnis der von der Flüssigkeit berührten zur ganzen Seitenwand das Verhältnis des den Austritt der Flüssigkeit bewirkenden Teilbereichs zum ganzen Bereich der allgemeinen Bedingungen.

Aber nicht alle Urteilsmaterien, welche zur Aufstellung disjunktiver Urteile Anlaß bieten, sind von der hier dargelegten Beschaffenheit. Wir wollen zwei Beispiele anderer Art vorführen und analysieren.

„Wenn ich in eine Urne mit undurchsichtigen Wänden, die eine weiße, zwei schwarze und drei rote, im übrigen gleiche Kugeln enthält, greife und eine Kugel hervor hole, so ist dieselbe weiß oder schwarz oder rot.“ Auch hier kann von einem Bereich der allgemeinen Bedingungen und von einer Scheidung desselben in Teilbereiche gesprochen werden, die den Disjunktionsgliedern zugeordnet sind, und zwar in folgendem Sinne. Bei einer bestimmten Anordnung der Kugeln in der Urne führt jede der unendlich vielen Bewegungen, welche die vollziehende Hand in Ausführung der allgemeinen Voraussetzung machen kann, mit Notwendigkeit zu einer Kugel von bestimmter Farbe; das gleiche gilt für jede der unendlich vielen denkbaren Lagerungen der Kugeln in der Urne. Mit Rücksichtnahme auf alle diese Lagerungen läßt sich also eine Gesamtheit von Handbewegungen denken, die zu einer der Kugeln überhaupt hinführen, und in dieser lassen sich Teilgesamtheiten unterscheiden, von welchen die eine mit Notwendigkeit die weiße, eine andere eine schwarze, die dritte eine rote Kugel zur Folge hat. Aber die Bewegungen, die zu dem einen oder andern Erfolge hinleiten, zu charakterisieren, dafür gibt es keinen Anhalt; daher ist es auch untunlich, das disjunktive Urteil in hypothetische aufzulösen, d. h. die allgemeine Voraussetzung so zu spezialisieren, daß sich ein bestimmter Erfolg mit Notwendig-

keit daran knüpfe. Dagegen gestattet die dargelegte allgemeine Erkenntnis die quantitative Vergleichung der verschiedenen Bereiche von Bewegungen. Denkt man sich nämlich innerhalb einer jeden *räumlichen* Lagerung der sechs Kugeln alle möglichen Anordnungen der Farben ausgeführt, so wird ein bestimmter Platz zweimal so oft durch eine schwarze und dreimal so oft durch eine rote besetzt sein als durch eine weiße: daher verhalten sich die Gesamtheiten der Bewegungen, welche zu einer weißen, einer schwarzen, einer roten Kugel führen, quantitativ wie $1:2:3$.

Ein von den bisherigen wesentlich verschiedenes Beispiel eines disjunktiven Urteils ist das folgende: „Wenn ein Mensch das Alter von x Jahren erreicht hat, so erlebt er das Alter von $x+1$ Jahren oder er erlebt es nicht.“ Unter einer Verwirklichung der allgemeinen Voraussetzung ist hier die Beobachtung eines Individuums zu verstehen, welches x Jahre alt geworden und nun mit der ihm eigentümlichen Konstitution den kaum dem Namen nach anführbaren Einflüssen ausgesetzt ist, von welchen Leben und Sterben abhängt und die in einer bestimmten Kombination auf dasselbe einwirken. Man kann sich nun ganz wohl vorstellen, daß sich in Ansehung einer bestimmten Konstitution die Kombinationen der maßgebenden Einflüsse in zwei Gebiete scheiden, die einen mit Notwendigkeit zum Durchleben des gedachten Zeitraumes, die andern zum vorzeitigen Tode führend. Über diese allgemeine Vorstellung ist aber mit dem zu Gebote stehenden Wissen über den Sachverhalt nicht hinauszukommen. Zu quantitativen Vergleichen und zur Bildung hypothetischer Urteile ist hier kein Anhalt gegeben.

3. Gegenstand der Wahrscheinlichkeitstheorie. Das disjunktive Urteil steht mit dem Gegenstande der Wahrscheinlichkeitstheorie in engem Zusammenhange, es kann als ihre logische Grundlage bezeichnet werden [1]). Das praktische Leben stellt uns oft vor Materien, welche die Bildung hypothetischer Urteile ausschließen, weil die allgemeinen Bedingungen mehrere einander ausschließende Erfolge *möglich* erscheinen lassen, ohne daß aus dem Gange der Verwirklichung jener Bedingungen jemals ein bestimmter Erfolg mit Sicherheit zu erschließen wäre.

Die vorgeführten Beispiele haben nun gezeigt, daß derlei Urteilsmaterien bei Vorhandensein eines bestimmten *Wissens* die Ausführung *quantitativer* Bestimmungen zulassen; dadurch aber können sie der mathematischen Behandlung zugänglich gemacht werden. Jene Bestimmungen beziehen sich auf den Gesamtumfang der Ausführungsmöglichkeiten der allgemeinen Bedingungen und auf jenen Teil des Gesamtumfangs, welcher mit Notwendigkeit zu einem bestimmten

1) Ch. Sigwart, Logik, II. Bd., 1893, p. 305 und 310, Fußnote.

Erfolge hinführt; ist es gelungen, diese Umfänge durch Zahlen zu charakterisieren, so kann der *Quotient* aus der zweiten Zahl durch die erste dazu dienen, die Stellung dieses Erfolges in der Gesamtheit der möglichen Erfolge zu kennzeichnen. Er soll, ohne daß zunächst auf seine etwaige innere Bedeutung eingegangen würde, als die *Wahrscheinlichkeit* des betreffenden Erfolges definiert werden.

Durch die Bestimmung der Wahrscheinlichkeit erfährt das ursprüngliche disjunktive Urteil eine wesentliche Ergänzung. Während sich dasselbe beschränkt, die möglichen Folgen einer Voraussetzung auseinander zu halten und aufzuzählen, wird nun jeder dieser Folgen eine Zahl koordiniert, wodurch die ganze Urteilsmaterie eine bestimmte Charakterisierung empfängt.

Der Sachverhalt läßt auch die folgende Auffassung zu[1]). Man kann aus einem hypothetisch-disjunktiven Urteil neue Urteile bilden, indem man die allgemeine Voraussetzung mit jedem einzelnen Gliede der Disjunktion verbindet. Ein solches Urteil[2]) hat wohl die *Form*, nicht aber das *Wesen* eines hypothetischen Urteils: während diesem das Prädikat der *Notwendigkeit* oder der *Gewißheit* (für den Fall der Realisierung der Voraussetzung) zukommt, kann bei jenem nur von *Möglichkeit* oder unter Umständen von *Wahrscheinlichkeit* gesprochen werden.

Es wird sich empfehlen, hier über einige üblichen Ausdrücke und Ausdrucksweisen sowie über einige den Wahrscheinlichkeitsbegriff in seiner bisherigen Fassung betreffenden Fragen ins klare zu kommen.

Wir haben bisher von den Gliedern der Disjunktion als von Folgen der Voraussetzung oder von möglichen Erfolgen gesprochen. Die Wahrscheinlichkeitstheorie bedient sich für einen Erfolg fast allgemein des Ausdrucks *Ereignis*, wiewohl derselbe nicht immer zutreffend ist; häufig würde sich die Bezeichnung *Tatbestand* besser empfehlen[3]). Denn daß eine Division ohne Rest aufgeht oder daß eine Gerade eine andere schneidet, ist nicht so sehr Ereignis als vielmehr ein Sachverhalt, ein Tatbestand. Indessen kann aus der Beibehaltung des in alle Literaturen übergegangenen Ausdrucks „Ereignis“ kein Mißverständnis entspringen.

Man schreibt allgemein die Wahrscheinlichkeit dem *Ereignis* oder Erfolg zu, wiewohl sich dagegen einwenden läßt, daß einem Ereignis an sich eine Wahrscheinlichkeit nicht zukommt, sondern nur im Hinblick auf die Bedingungen, daß also nur von der Wahrscheinlichkeit eines *Urteils* gesprochen werden könne. Dieser Standpunkt ist von

---

1) A. Fick, Philosophischer Versuch über die Wahrscheinlichkeiten, 1838.

2) Von Fick als „unvollständig ausgedrücktes hypothetisches Urteil“ bezeichnet l. c. I, p. 12.

3) K. Stumpf, Über den Begriff der mathematischen Wahrscheinlichkeit. Ber. d. bayr. Akad. (phil. Kl.), 1892, p. 46.

Fick[1]) ausdrücklich vertreten worden, der nur die Wahrscheinlichkeit eines unvollständig ausgedrückten hypothetischen Urteils gelten lassen will, denn „jedes individuelle Ereignis ist entweder im Kausalnexus notwendig und wirklich oder unmöglich". Gegen die letztere Bemerkung ist sicherlich nichts einzuwenden, aber ein „individuelles" Ereignis kann niemals den Gegenstand der Wahrscheinlichkeitstheorie bilden, sondern immer nur *das* Ereignis in abstracto. Um den Unterschied der beiden Ausdrucksweisen deutlich zu machen, möge ein Beispiel angeführt werden. Nach Fick kommt dem unvollständig ausgedrückten hypothetischen Urteil: „Wenn man eine dekadische Zahl durch 5 dividiert, so geht die Division auf" — die Wahrscheinlichkeit $\frac{1}{5}$ zu; der allgemeine und nach unserem Dafürhalten richtige Sprachgebrauch schreibt sie dem Aufgehen der Division zu; dem steht keineswegs entgegen, daß eine individuelle Ausführung der Voraussetzung entweder sicher ohne Rest oder mit einem solchen abläuft.

Wenn wir zu dem Beispiel der Urne mit der weißen, den schwarzen und roten Kugeln zurückkehren, so gewahren wir ein *positives Wissen*, welches bei der quantitativen Vergleichung der Gesamtheiten von Bewegungen, die zu bestimmten Erfolgen hinführen, maßgebend war: es ist dies die Kenntnis der Anzahl der Kugeln jeder Farbe. Im Falle der Realisierung treten uns Umstände entgegen, die sich unserer Kenntnis entziehen und von einer Realisierung zur andern in einer der Erkenntnis unzugänglichen Weise ändern: die Lagerung der Kugeln in der Urne und die Bewegung der ziehenden Hand. Jenes positive Wissen ist das *Bleibende*, das *Konstante* an der Urteilsmaterie und verleiht der aus ihr abgeleiteten Wahrscheinlichkeit *objektive*, weil vom urteilenden Subjekt unabhängige *Bedeutung*. Gegenüber einer Urteilsmaterie, bei der ein so geartetes positives Wissen fehlt, kann ein Subjekt auf Grund des ihm zu Gebote stehenden allgemeinen Wissens an eine Schätzung der Umfänge der Realisierungsmöglichkeiten der einzelnen Disjunktionsglieder gehen und aus den Resultaten dieser Schätzung Wahrscheinlichkeiten bilden; diese haben jedoch nur *subjektive Bedeutung*, können sich von Subjekt zu Subjekt und selbst bei demselben Urteilenden ändern, wenn sein Wissen von der Materie einen andern Umfang annimmt.

Ist aus dem Gange der Erwägungen, welche uns zu dem Wahrscheinlichkeitsbegriff geführt haben, die Erkenntnis erwachsen, daß die Wahrscheinlichkeit eine der Urteilsmaterie als solcher anhaftende Eigenschaft, eine Funktion dieser Materie sei, so folgt daraus von selbst, daß sie an jedem Ort und zu jeder Zeit dieselbe bleibt, daß also insbesondere die Zeit der Verwirklichung eines Erfolges ohne

---

1) l. c. p. 12 u. 14.

Einfluß ist auf seine Wahrscheinlichkeit. Man ist wohl gewohnt, mit dem Begriff Wahrscheinlichkeit die Vorstellung eines *zukünftigen* Erfolges zu verknüpfen, und Lotze[1]) war der Meinung, daß nur diese Verbindung einen Sinn habe; mit derselben Berechtigung kann aber von der Wahrscheinlichkeit eines bereits realisierten, aber unbekannt gebliebenen Erfolges gesprochen werden, wiewohl er schon wirklich vorhanden ist oder war. Die beiden Fragen: „Aus der oben gedachten Urne wird eine Kugel gezogen werden; mit welcher Wahrscheinlichkeit wird die gezogene Kugel weiß sein?“ und: „Aus der Urne ist (zu welcher Zeit immer) eine Kugel gezogen worden; welches ist die Wahrscheinlichkeit, daß sie weiß war?“ haben logisch ganz gleiche Bedeutung und sind auch in gleicher Weise zu beantworten.

Der vorstehenden Erörterung lag die Annahme zugrunde, daß das Substrat des positiven Wissens, welches uns bezüglich der Materie zu Gebote steht und das bei der Wahrscheinlichkeitsbestimmung maßgebend ist, im Laufe der Zeit unverändert, konstant bleibe. Es gibt aber auch Urteilsmaterien, bei welchen jenes Substrat einer zeitlichen Änderung unterworfen ist; wenn solche Materien eine Wahrscheinlichkeitsbestimmung zulassen, so wird die *Wahrscheinlichkeit eine Funktion der Zeit* sein. Besteht z. B. das Substrat in physischen Körpern, welche ihren Zustand zeitlich ändern, und ist die Wahrscheinlichkeitsbestimmung von dem Zustand abhängig, so wird auch sie einer zeitlichen Änderung unterworfen sein. In einem früheren Beispiel ist von den beiden Möglichkeiten die Rede gewesen, welchen eine Person, die das Alter x erreicht hat, während eines Jahres ausgesetzt ist; angenommen, es wäre möglich, die Wahrscheinlichkeit zu bestimmen, daß sie dieses Jahr durchlebe, so ist anzunehmen, daß diese Wahrscheinlichkeit für die gleiche Altersstufe zu einer andern Zeit eine andere sein werde; denn die Umstände, welche auf das Leben und Sterben zweifellos Einfluß üben (wie z. B. das Klima, die wirtschaftlichen Verhältnisse, die hygienischen Einrichtungen), sind zeitlicher Änderung unterworfen. Die im vorigen Absatz aufgestellte Behauptung, daß der Wahrscheinlichkeitsbegriff mit zukünftigen wie mit vergangenen Realisierungen in Verbindung gebracht werden könne, bleibt von diesen Erwägungen unberührt.

4. Der Zufall. Bei der Analyse des Urnenbeispiels haben wir neben einem positiven Wissen Umstände erkannt, die bei der Realisierung der allgemeinen Bedingungen, d. i. bei Ausführung eines Zuges, auftretend sich einzeln und im Zusammenwirken unserer Kenntnis entziehen und uns verhindern, die Notwendigkeit für das Eintreten oder für das Ausbleiben eines bestimmten Erfolges zu erkennen. Wenn wir den Erfolg oder das Ereignis, auf welches unser Interesse sich

1) R. H. Lotze, Logik, 1893.

richtet, mit Rücksicht auf diese Umstände als ein *zufälliges* bezeichnen, so negieren wir damit *die Notwendigkeit* seines Eintreffens; und wenn wir sagen, sein Eintreffen hänge vom *Zufall* ab, so meinen wir hierunter das Zustandekommen einer solchen Verbindung der voneinander unabhängigen Umstände, welche das Ereignis wirklich herbeiführt. Ist beispielsweise unser Interesse dem Erscheinen der weißen Kugel zugewandt, so bezeichnen wir es als zufällig, daß bei der uns unbekannten Lagerung der Kugeln und der davon unabhängigen Bewegung der Hand diese gerade die weiße Kugel erfassen sollte, und wenn wir ein solches Geschehen als durch den Zufall hervorgebracht erklären, so fassen wir den *Zufall* als ein *gesetzlos Wirkendes* im Gegensatze zu einem *gesetzmäßig Wirkenden*, einer *Ursache*, auf; aus der gesetzlosen Wirksamkeit des Zufalls erklären wir auch die *Unregelmäßigkeit* in einer Reihe von Erfolgen.

Indessen ist die Vorstellung von dem Walten eines „blinden Zufalls" unvereinbar mit unserer Überzeugung von dem kausalen Zusammenhange alles Geschehens, eine bloße Fiktion. Jedes Geschehen ist die Folge eines vorausgegangenen andern und die Veranlassung eines künftigen, und so ist alles Geschehen nach dem Prinzip der Ursächlichkeit untereinander verbunden. Wenn eine Person wiederholt würfelt, so ist die Endlage eines Wurfs eine notwendige Folge der Endlage des vorigen Wurfs und der inzwischen dem Würfel erteilten Impulse und Bewegungen; für eine Intelligenz, die das alles zu durchschauen und nach seinen Wirkungen zu beurteilen vermöchte, wäre das Resultat nichts Ungewisses, Unbekanntes. Selbst bei den Würfen, die zwei verschiedene Personen ausführen, kann man nicht in aller Strenge von Unabhängigkeit reden, weil in der Art und Weise der menschlichen Bewegungen und den sonstigen Anordnungen etwas Gemeinsames gelegen ist.

Mit dem Worte Zufall soll also der kausale Zusammenhang nicht geleugnet, durch seinen Gebrauch vielmehr das Zugeständnis gemacht werden, daß uns der Zusammenhang in dem betreffenden Falle überhaupt nicht erkennbar (Auftreten einer Epidemie und Sichtbarkeit eines Kometen) oder so verwickelt sei (wie etwa beim Würfeln), daß wir über seine bloße Konstatierung nicht hinauskommen.

Es ist üblich, die in das Gebiet der Wahrscheinlichkeitstheorie fallenden Ereignisse, Erscheinungen, Tatbestände dadurch allgemein zu charakterisieren, daß ihr Eintreffen oder ihre Existenz einerseits von bekannten und bleibenden und andererseits von unbekannten und wechselnden Umständen abhänge. Die Umstände der ersten Art bezeichnet man als *Ursachen*, die der zweiten Art als *Zufall*.

Man bezeichnet die Wahrscheinlichkeitsrechnung auch als „Theorie des Zufalls" und will damit einen Teil ihres Anwendungsgebiets kennzeichnen, solche Erscheinungen nämlich, bei denen wir den ursäch-

lichen Zusammenhang mit andern nicht anzugeben vermögen; nicht aber soll mit jener Bezeichnung gemeint sein, daß es sich um die Erforschung des Unregelmäßigen, des Gesetzlosen, um es kurz zu sagen, des Zufälligen im Verlaufe solcher Erscheinungen handelt.

5. Gleichmögliche Fälle. Die Materien, auf welche sich die ersten Probleme der Wahrscheinlichkeitsrechnung bezogen, hatten ihren Ursprung in *Glücksspielen*: in Würfel- und Kartenspielen, in Ziehungen von Nummern oder Kugeln verschiedener Farben aus Urnen u. dgl. Solche Materien sind für die Bildung von Wahrscheinlichkeiten typisch geworden, und zwar aus dem Grunde, weil hier das positive Wissen eine bestimmte leicht zu überblickende Struktur besitzt. Es läßt sich nämlich die Disjunktion bis zu sogenannten *möglichen Fällen* hinführen, die zunächst nur in dem Sinne gleichartig sind, als sie eine weitere Disjunktion nicht zulassen und daher die einfachsten Glieder des disjunktiven Urteils darstellen. Nehmen wir das Beispiel mit der Urne wieder auf, welche eine weiße, zwei schwarze und drei rote Kugeln enthält, so läßt sich zunächst eine Disjunktion nach den Farben bilden, deren Glieder aber insofern nicht gleichartig sind, als sich zwei davon weiter auflösen lassen, wenn man die Kugeln einer Farbe individualisiert; nach dieser Auflösung entstehen sechs mögliche Fälle, dargestellt durch die einzelnen Kugeln. Der Würfel, der auf den Tisch hingeworfen wird, bietet sechs mögliche Fälle dar, vertreten durch die mit verschiedenen Punktzahlen bezeichneten Seitenflächen. In einem Kartenspiel bilden die einzelnen Blätter die möglichen Fälle.

Soll nun auf diese Disjunktion möglicher Fälle eine Wahrscheinlichkeitsbestimmung gegründet werden, so muß bekannt sein, welcher Teil des Gesamtbereichs der Ausführungsmodalitäten jedem einzelnen Falle zugeordnet ist in dem Sinne, daß die spezielle Ausführung der allgemeinen Bedingungen, wenn sie in einen solchen Teil fällt, den betreffenden Fall notwendig herbeiführt. Die einfachste Teilung des Gesamtbereiches ist nun die, daß den einzelnen Fällen gleiche Teile desselben zukommen. Die einzelnen Fälle heißen dann *gleichmögliche Fälle*; in jüngster Zeit sind dafür auch die Bezeichnungen „gleichberechtigte“[1]) und „gleichwertige“[2]) Fälle, Annahmen u. dgl. gebraucht worden.

Eine für die Wahrscheinlichkeitstheorie fundamentale Frage geht nun dahin, auf welcher Grundlage über die Gleichmöglichkeit der Fälle entschieden werden soll. Um zur Beurteilung dieser Frage Stellung nehmen zu können, gehen wir zunächst auf die Erklärungen zurück, welche zwei Klassiker unseres Gegenstandes, Jacob Bernoulli und Laplace, für die gleichmöglichen Fälle gegeben haben.

---

1) J. v. Kries, Die Prinzipien der Wahrscheinlichkeitsrechnung, 1886.
2) Ch. Sigwart, Logik, II, p. 305.

Bernoulli[1]) erklärt die sechs Fälle, welche ein Würfel darbieten kann, für gleichmöglich, da „wegen der gleichen Gestalt aller Flächen und wegen des gleichmäßig verteilten Gewichtes des Würfels *kein Grund* dafür vorhanden ist, daß eine Würfelseite leichter als eine andere fallen sollte". Laplace[2]) umschreibt den Terminus „gleichmögliche Fälle" das eine Mal mit den Worten, es seien Fälle, „über deren Dasein wir in *gleicher Unwissenheit* sind, und erklärt sie ein zweites Mal damit, man „habe *keinen Grund* zu glauben, einer der Fälle werde eher eintreten als die andern".

Beide Erklärungen berufen sich auf das logische *Prinzip des mangelnden Grundes*, das hier zu folgendem Schlusse benützt wird: Weil kein Grund erkennbar ist, der für eine ungleiche Realisierungsmöglichkeit der Fälle sprechen würde, so werden diese als gleichmöglich erachtet. Indessen ist nicht zu übersehen, daß Bernoulli in dem angeführten Beispiel wie auch an andern Stellen des zitierten Werkes nicht dieses Prinzip allein anwendet, sondern auch *für* die Gleichheit der Fälle Gründe anführt.

Ein Prinzip, welches dem vorigen gegenüber den entgegengesetzten Standpunkt einnimmt, ist das des *zwingenden Grundes*, welches die Gleichmöglichkeit der Fälle nur dann auszusprechen gestattet, wenn dafür entscheidende, zwingende Gründe vorhanden sind.

Diese beiden Prinzipe bezeichnen gewissermaßen die beiden extremen Standpunkte, auf welche man sich bei der Beantwortung der oben formulierten Frage stellen kann; von der Wahl zwischen beiden hängt die Ausdehnung des Anwendungsgebietes der Wahrscheinlichkeitsrechnung ab.

Zur Erläuterung mögen einige Beispiele dienen.

Weiß man von einer Urne bloß, daß sie nur weiße und schwarze Kugeln enthält, so kennt man wohl die Disjunktion „weiß oder schwarz", über ihre Glieder herrscht aber im übrigen absolutes Nichtwissen, was im Sinne der Laplaceschen Umschreibung als gleiches Nichtwissen ausgelegt werden kann, es besteht also auch kein Grund anzunehmen, daß der weißen Kugeln mehr oder weniger vorhanden seien als der schwarzen; folglich ist, wenn man den ersten Standpunkt einnimmt, die Wahrscheinlichkeit eines jeden Disjunktionsgliedes mit $\frac{1}{2}$ anzusetzen. Der zweite Standpunkt gestattet diesen Ansatz nur dann, wenn das positive Wissen vorhanden ist, daß von jeder Farbe gleichviel Kugeln in der Urne sich befinden.

Wird ein von sechs Flächen, welche durch aufgeschriebene Nummern voneinander unterschieden sind, begrenzter Körper hin-

---

1) Ars conjectandi, 1713, p. 224 (Ostwalds Klassiker, Nr. 108, p. 88).

2) Théorie analytique des probabilités (1812) p. 178—179.

geworfen, so besteht über die möglichen Fälle absolutes Nichtwissen, und es ist nach dem Grundsatz des mangelnden Grundes jedem die Wahrscheinlichkeit $\frac{1}{6}$ zuzuschreiben. Das Prinzip des zwingenden Grundes verlangt aber ein ganz bestimmtes positives Wissen über die physische Konstitution des „Körpers", wenn dieser Ansatz zulässig sein soll: es setzt die Würfelform und eine derartige Verteilung der Masse voraus, daß der Schwerpunkt mit dem geometrischen Mittelpunkt zusammenfällt.

Der Frage gegenüber, welche Wahrscheinlichkeit es hat, daß ein Himmelskörper, z. B. der Sirius, Eisen enthalte, antwortet das erste Prinzip, auf das völlige Nichtwissen über das „ja" und „nein" gestützt, mit $\frac{1}{2}$. Das andere Prinzip wird es, eben mangels jeglichen positiven Wissens, ablehnen, eine Antwort zu geben. Ganz allgemein: Aus dem bloßen Urteile: „Wenn A, so ist B oder nicht B" wird auf der einen Seite für die Aussage: „Wenn A, so ist B" immer auf die Wahrscheinlichkeit $\frac{1}{2}$ geschlossen, während die andere Seite ein solches Urteil nicht als Gegenstand der Wahrscheinlichkeitsrechnung zuläßt.

Diese Ausführungen sind wohl geeignet, dem zweiten Standpunkt den Vorzug zu geben, soll anders den Wahrscheinlichkeiten nicht bloß eine formale, sondern eine objektive Bedeutung zukommen, sollen sie der Ausdruck eines bestimmten Wissens und nicht ein bloßes Symbol für das Nichtwissen über eine Urteilsmaterie sein.

Indessen soll nicht verschwiegen werden, daß es in den seltensten Fällen gelingen wird, die Gleichmöglichkeit in jener Strenge zu erweisen, wie es das Prinzip des zwingenden Grundes fordert. Man wird bezüglich einer Münze, eines Würfels, wenn sie noch so genau den an sie zu stellenden physischen Anforderungen entsprechen, wohl nie mit apodiktischer Gewißheit sagen können, daß für die einzelnen Seiten gleiche Realisierungsmöglichkeit bestehe. Wenn man trotzdem von gleichmöglichen Fällen spricht, so stützt man sich dabei bezüglich des Restes doch wieder auf das Prinzip des mangelnden Grundes: weil man keinen Grund hat, anzunehmen, daß die etwa vorhandene Abweichung gerade diese oder jene Seite begünstige, schreibt man den Seiten gleiche Realisierbarkeit zu. Daß dies aber ein ganz anderer Schluß ist als der bei dem „sechsflächigen Körper" gemachte, ist leicht zu erkennen; bei dem Spielwürfel, wenn er nicht auffällig von dem Typus des Würfels abweicht, *wird* die (wohl sicher vorhandene) Ungleichheit bezüglich der einzelnen Seiten keine erhebliche sein; bei jenem Körper von unbekannter Form und Massenverteilung *kann* sie aber sehr beträchtlich sein. Man kann im allgemeinen sagen, daß die über konkrete Urteilsmaterien aufgestellten Wahrscheinlich-

keiten mit den Ergebnissen einer materiellen Messung das gemein haben, innerhalb gewisser Grenzen unsicher zn sein; die jeweilige Weite dieser Grenzen (sofern es denkbar wäre, sie zu ermitteln) wäre ein Maß für die Sicherheit, Genauigkeit oder für den *Erkenntniswert*[1]) der betreffenden Wahrscheinlichkeit.

Für die Beurteilung des Möglichkeitsgrades der einzelnen Fälle ist jedoch das feste positive Wissen, welches man über dieselben hat, nicht allein maßgebend; es kommt vielmehr auch auf den Vorgang bei der Verwirklichung der allgemeinen Bedingungen an. Dieser muß so geartet sein, daß er jede Betätigung einer Absicht desjenigen, der die Ausführung vornimmt, auf die Begünstigung eines oder einer Gruppe von Fällen vollkommen ausschließt. Bei wiederholten Realisierungen muß, wenn dies so ausgedrückt werden darf, alles vorgekehrt werden, was geeignet ist, allen einzelnen Fällen gleiche Gelegenheit zu ihrer Realisierung zu gewähren. Wenn jemand mit einem „vollkommenen“ Würfel einen „Wurf“ derart ausführt, daß er dem Würfel erst eine bestimmte Lage gibt und ihn dann aus geringer Höhe auf den Tisch frei fallen läßt, so schließt dieser Vorgang die Begünstigung des Erscheinens einer bestimmten Seite nicht aus, und man kann einer solchen Verwirklichung des „Werfens“ gegenüber trotz der (vorausgesetzten) Vollkommenheit des Würfels den einzelnen Seiten nicht mehr gleichen Möglichkeitsgrad zusprechen. Hiernach ist wohl zu erkennen, warum man bei dem Spiel „Wappen—Schrift“ die Münze hoch emporschleudert, bei dem Würfelspiel den Würfel erst in einem Becher schüttelt und dann hinwirft, die Blätter eines Kartenspiels auf der Rückseite so gleichartig als möglich gestaltet und sie mischt, die Nummern oder Kugeln in einer Urne zwischen aufeinanderfolgenden Ziehungen durcheinander mengt usw.

Die möglichen Fälle einer Urteilsmaterie scheiden sich in bezug auf einen bestimmten Erfolg oder auf ein Ereignis in solche, mit deren Verwirklichung der Eintritt des Erfolges verbunden ist — die *günstigen* Fälle — und in die übrigen, welche als die dem betreffenden Ereignis *ungünstigen* bezeichnet werden. Bei dem Urnenbeispiel (eine weiße, zwei schwarze, drei rote Kugeln) sind dem Erscheinen einer schwarzen Kugel zwei Fälle günstig und vier Eälle ungünstig.

Mitunter gebraucht man für die günstigen Fälle auch die Bezeichnung *Chancen* des Ereignisses. Indessen wird dies Wort in der Wahrscheinlichkeitstheorie mit wechselnder Bedeutung angewandt; so werden zuweilen die bleibenden bekannten Umstände, für welche im vorigen Artikel der Name *Ursachen* (der möglichen Erfolge) angeführt worden ist, auch Chancen genannt; nicht selten wird das Wort *Chance*

1) Man vergleiche hierzu K. Stumpf, l. c., und die Kritik A. Meinongs über v. Kries' Prinzipien der Wahrscheinlichkeitsrechnung, Gött. gel. Anz. 1890.

als Synonym für Wahrscheinlichkeit (insbesondere dann, wenn subjektive Momente ausgeschlossen sind) gebraucht.

6. Die Voraussetzungen der Wahrscheinlichkeitsrechnung. Es wird sich empfehlen, am Schlusse dieser grundlegenden Betrachtungen das Operationsfeld der Wahrscheinlichkeitsrechnung näher zu kennzeichnen.

Die Wahrscheinlichkeitstheorie abstrahiert von dem kausalen Zusammenhange des Geschehens, stellt also die Hypothese *rein zufälliger* Ereignisse auf, womit zugleich die völlige gegenseitige *Unabhängigkeit* der Vorgänge supponiert ist, welche solche Ereignisse herbeiführen. Sie nimmt ferner bei den möglichen Fällen, mit welchen sie operiert, *Gleichmöglichkeit* an.

Den aus diesen Prämissen durch logische Schlüsse abgeleiteten Sätzen der Wahrscheinlichkeitsrechnung kommt ebenso wie den Lehrsätzen der andern Teile der reinen Mathematik Gewißheit zu. Es entsteht aber die Frage, ob ihnen auch irgendeine Bedeutung für die Wirklichkeit zukommt, da in dieser die gemachten Voraussetzungen niemals streng erfüllt sind.

Auf diese Frage ist zunächst mit dem Hinweis auf andere Gebiete der Wissenschaft zu erwidern. Die Geometrie geht beispielsweise auch von Grundbegriffen aus, die nirgends vollkommen verwirklicht sind, wie Punkt, Gerade, Ebene; von ihren Sätzen macht man aber trotzdem mit Erfolg Gebrauch bei Untersuchungen über die räumliche Anordnung der Materie. In jenen Wissenschaften, die auch mit den inneren Eigenschaften der Materie zu rechnen haben, geht die Abstraktion viel weiter; die Vorstellungen von der stetigen Erfüllung des Raumes durch Materie, von starren, vollkommen unelastischen, von vollkommen elastischen Körpern, von gewichtlosen, biegsamen und nicht dehnbaren Fäden u. ä., deren man sich in der Mechanik bedient, um Probleme der mathematischen Behandlung zugänglich zu machen, die ohne solche Abstraktionen gar nicht in Angriff genommen werden könnten, sind Hypothesen, die sich nirgends streng verwirklicht finden. Und doch lehrt die tägliche Erfahrung, daß sich mit den Sätzen der Mechanik Probleme der Praxis in brauchbarer Weise behandeln lassen. Die Erfahrung ist auch das einzige Mittel, die Anwendbarkeit theoretischer Sätze auf die Wirklichkeit zu erproben; auf rein logischem Wege läßt sich darüber keine Entscheidung treffen.

Es ist also durchaus nicht ausgeschlossen, daß auch die Sätze der Wahrscheinlichkeitsrechnung sich werden auf wirkliche Vorgänge anwenden lassen. Freilich könnte ein Zweifel in der Richtung bestehen bleiben, daß, um bei dem Beispiel der Geometrie zu verbleiben, die Grundbegriffe hier durch Abstraktion aus der Wirklichkeit gewonnen sind, daß es also nicht überraschen könne, wenn die Sätze

sich in der Wirklichkeit mit größerer oder geringerer Annäherung bestätigt finden; wogegen die Grundbegriffe der Wahrscheinlichkeitstheorie anscheinend freie Schöpfungen des Geistes sind. Dem ist aber in der Tat nicht so; auch diese Begriffe sind realen Ursprungs, nämlich aus dem Bestreben hervorgegangen, die relative Häufigkeit des Auftretens bestimmter Spielergebnisse aus den Bedingungen des Spiels zu erklären.

7. Definition der mathematischen Wahrscheinlichkeit. Wenn das positive Wissen über eine Urteilsmaterie die Auflösung der Möglichkeiten in eine zählbare Menge einzelner gleichmöglicher Fälle gestattet, *so drückt sich die Wahrscheinlichkeit eines* auf der Urteilsmaterie beruhenden *Ereignisses als Quotient aus der Anzahl der ihm günstigen durch die Anzahl der gleichmöglichen Fälle aus.*

Durch die der Definition vorangestellten Worte ist schon angedeutet, daß sie zur Erledigung nur eines bestimmten Teils von Fragen ausreichen werde. Die Definition wird daher im Verlaufe der Untersuchungen Erweiterung und Ausgestaltung erfahren müssen.

Indem wir ein Ereignis mit E bezeichnen, soll seine Wahrscheinlichkeit durch das Symbol $\mathfrak{W}(E)$ ausgedrückt werden. Die Form der Funktionsbezeichnung erinnert daran, daß die Urteilsmaterie dem Ereignis eine bestimmte Wahrscheinlichkeit *zuordnet.*

Ist also g die Anzahl der dem Ereignis E günstigen und m die Anzahl aller möglichen Fälle, so ist

$$\mathfrak{W}(E) = \frac{g}{m} \tag{1}$$

die Wahrscheinlichkeit von E.

Eine solche zahlenmäßig bestimmte Wahrscheinlichkeit nennt man zum Unterschiede von dem allgemeinen Begriff der Wahrscheinlichkeit als Gegensatz der Notwendigkeit wohl auch eine *mathematische Wahrscheinlichkeit.*

Die Methode, welche in der obigen Definition ihren Ausdruck findet, bezeichnet man als eine *Wahrscheinlichkeitsbestimmung a priori.* Mitunter wird diese Bezeichnung[1]) auch auf die Wahrscheinlichkeit selbst übertragen und dann von einer „Wahrscheinlichkeit a priori" oder von einer „apriorischen Wahrscheinlichkeit" gesprochen.

Das Zahlengebiet, auf welchem sich in solcher Weise bestimmte Wahrscheinlichkeiten bewegen, ist das der (positiven) echten Brüche, also der *rationalen* Zahlen aus dem Intervall (0, 1) mit Ausschluß der Grenzen; denn diese Grenzen charakterisieren eine *Notwendigkeit*

---

1) Dieselbe ist in diesem Zusammenhange zuerst von Jacob Bernoulli, Ars conjectandi, p. 224 (Ostwalds Klassiker Nr. 108, p. 89), gebraucht worden. Über die zweite Methode, von ihm „Wahrscheinlichkeitsbestimmung a posteriori" genannt, handelt der III. Abschnitt.

oder *Gewißheit*; die untere, aus $g = 0$ entspringend, drückt die Notwendigkeit des Nichteintreffens, die obere, aus $g = m$ hervorgehend, die Notwendigkeit des Eintreffens von E aus. Eine zutreffende geometrische Darstellung der Wahrscheinlichkeiten und der Notwendigkeit böte eine Kreislinie vom Umfange 1 dar: ein Punkt derselben wäre das Bild der Notwendigkeit, zur einen Seite (0) des Nichteintreffens, zur andern Seite (1) des Eintreffens.

Die $g' = m - g$ Fälle, welche nach Ausscheidung der günstigen verbleiben, nennt man in bezug auf das Ereignis E *ungünstige* Fälle und die Verwirklichung eines derselben ein dem E entgegengesetztes Ereignis $\bar{E}$ (gleichbedeutend mit dem Nichteintreffen von E). Seine Wahrscheinlichkeit $\mathfrak{W}(\bar{E})$ ist durch den Bruch $\frac{g'}{m}$ gegeben, also

$$\mathfrak{W}(\bar{E}) = \frac{g'}{m} = \frac{m-g}{m} = 1 - \mathfrak{W}(E),$$

woraus

$$\mathfrak{W}(E) + \mathfrak{W}(\bar{E}) = 1 \tag{2}$$

folgt, so daß zwei entgegengesetzte Ereignisse Wahrscheinlichkeiten besitzen, die sich zur Einheit ergänzen. Von dieser Beziehung wird zuweilen vorteilhafter Gebrauch gemacht, wenn die Berechnung von $\mathfrak{W}(\bar{E})$ sich leichter gestaltet als die des fraglichen $\mathfrak{W}(E)$.

Sind $E_1, E_2 \cdots E_n$ Ereignisse, welche im Bereiche der Möglichkeit liegen und einander ausschließen; entsprechen unter den m möglichen Fällen diesen Ereignissen beziehungsweise $g_1, g_2, \cdots g_n$ günstige Fälle, so folgt aus $g_1 + g_2 + \cdots + g_n = m$, daß

$$\mathfrak{W}(E_1) + \mathfrak{W}(E_2) + \cdots + \mathfrak{W}(E_n) = 1. \tag{3}$$

Diese Gleichung ist eine Erweiterung der Gleichung (2) und besagt, daß die Wahrscheinlichkeiten einer Reihe einander ausschließender Ereignisse die Summe 1 ergeben.

8. Bedeutung der Wahrscheinlichkeit. Wenn sie es auch nicht deutlich ausgesprochen, so haben doch schon die ersten Schriftsteller auf unserem Gebiete unter der Wahrscheinlichkeit ein *Maß für die Berechtigung einer Erwartung* oder *Vermutung* verstanden. Jacob Bernoulli hat diese Auffassung schon durch den Titel zum Ausdruck gebracht, den er seinem berühmten Werke gab: Ars conjectandi, d. i. Vermutungskunst. Die neuere philosophische Forschung hat sich dieser Deutung angeschlossen. v. Kries[1]) erkennt es als charakteristische Eigentümlichkeit einer Wahrscheinlichkeitsaussage, daß sie die mehr oder minder große Berechtigung einer Erwartung angebe, wenn auch jedesmal die Nichterfüllung derselben als möglich erscheint. Stumpf[2]) weicht hiervon nur insofern ab, als er von einer *vernünf-*

1) l. c., p. 21 und 285. 2) l. c.

tigen Erwartung gesprochen wissen will, womit gesagt sein soll, daß bei der Bildung der numerischen Wahrscheinlichkeit nur Vernunftgründe und nicht auch seelische Affekte zu Worte kommen sollen; wie man erkennt, kann diese Bemerkung sich nur auf *subjektive* Wahrscheinlichkeitsschätzungen beziehen. Das Wort *Erwartung* ersetzt er durch *Vermutung*, um der Vorstellung vorzubeugen, als ob es sich immer nur um Zukünftiges handelte.

Einen materiellen Ausdruck findet die Auffassung der Wahrscheinlichkeit als eines Erwartungsmaßes in der seit jeher üblichen Darstellung des Wahrscheinlichkeitsverhältnisses in Form einer *Wette*. Von dem Grundsatze ausgehend, die Einsätze in einer Wette müßten proportional sein der Stärke der Erwartung, daß das von dem einzelnen an der Wette Beteiligten bezeichnete Ereignis eintreffen werde oder eingetroffen sei, hat man die Wahrscheinlichkeit eines Ereignisses, dem g Fälle günstig und g' Fälle zuwider sind, mit den Worten gekennzeichnet: es sei g gegen g' auf das Eintreffen des Ereignisses zu wetten.

Nicht immer richtig ist das Verhältnis zwischen der verschiedener Grade fähigen Wahrscheinlichkeit und der absoluten Notwendigkeit oder Gewißheit beurteilt worden. Jacob Bernoulli[1]) bezeichnet die Wahrscheinlichkeit als einen Grad der Gewißheit und schreibt einem Ereignis von der Wahrscheinlichkeit $\frac{3}{5}$ den entsprechenden Bruchteil der Gewißheit zu; den gleichen Standpunkt hat auch noch der erste deutsche Philosoph, der sich mit der Kritik der Prinzipien der Wahrscheinlichkeitsrechnung befaßte, J. J. Fries[2]), eingenommen. Auch bei Laplace[3]) findet sich eine Stelle, die hieran anklingt; an die Bemerkung, daß die Wahrscheinlichkeit sich in Gewißheit verwandle und ihr Ausdruck die Einheit werde, wenn alle Fälle dem Ereignis günstig sind, knüpft er die Worte an, daß unter diesem Gesichtspunkte Gewißheit und Wahrscheinlichkeit vergleichbar seien; durch den weiteren Zusatz aber, daß ein wesentlicher Unterschied zwischen den beiden Zuständen des Geistes bestehe, wenn ihm eine Wahrheit streng bewiesen ist, oder wenn er noch eine kleine Quelle des Irrtums wahrnimmt, läßt er schon den richtigen Standpunkt durchblicken, auf den vor ihm schon Condorcet[4]) sich gestellt hat und den die neuere Philosophie behauptet: *Wahrscheinlichkeit und Gewißheit* (oder Notwendigkeit) *sind Dinge wesentlich verschiedener Natur*, und es gibt keine Brücke, die von der einen zur andern geschlagen werden

---

1) Ars conjectandi, p. 211 (Ostwalds Klassiker Nr. 108, p. 72).

2) Versuch einer Kritik der Prinzipien der Wahrscheinlichkeitsrechnung, 1842, p. 15.

3) Théorie analyt. d. probab. (1812) p. 179.

4) Essai sur l'application de l'analyse à la prob. etc. 1785.

könnte.[1]) Damit sind auch die mannigfachen Versuche ihrem logischen Werte nach gekennzeichnet, welche unternommen worden sind, um Zwischenglieder oder Übergänge zwischen Wahrscheinlichkeit und Gewißheit einerseits und Unmöglichkeit andererseits herzustellen. So unterscheidet schon Jacob Bernoulli[2]) zwischen *absoluter* und *moralischer* Gewißheit und Unmöglichkeit, unter letzteren sehr hohe, beziehungsweise sehr niedrige Grade von Wahrscheinlichkeit verstehend. Ähnliche Begriffsbildungen, welche gegen die *fundamentale* Erkenntnis verstoßen, daß ein Ereignis von einer der Einheit noch so nahen Wahrscheinlichkeit nicht eintreffen *muß* und ein Ereignis von noch so kleiner Wahrscheinlichkeit eintreffen *kann*, finden sich später bei D'Alembert[3]), Buffon[4]), De Morgan[5]) unter den Namen praktische Gewißheit, physische Unmöglichkeit u. dgl.

9. Literatur. Die ersten Probleme der Wahrscheinlichkeitsrechnung betrafen Glücksspiele und bezogen sich auf die mathematische Seite des Gegenstandes. Auf seine philosophische Seite ist Jacob Bernoulli im vierten Teile seiner „Ars conjectandi" 1713 (Ostwalds Klassiker Nr. 107 u. 108) zuerst des näheren eingegangen und hat zugleich zu zeigen sich bestrebt, daß die Anwendbarkeit und Bedeutung der Wahrscheinlichkeitsrechnung nicht bei den Glücksspielen aufhöre, sondern tief eingreife in viele Gebiete des praktischen Lebens. Laplace, dessen „Théorie analytique des probabilités" (1812, 1814, 1820; ges. Werke VII, 1886) ein Hauptwerk über unseren Gegenstand bildet und zugleich einen gewissen Abschluß seiner theoretischen Grundlegung bedeutet, hat sich in erster Linie der Ausbildung der mathematischen Seite und den Anwendungen zugewendet und ist weder hier noch in der ausdrücklich als „Essai philosophique des probabilités" bezeichneten Schrift, welche dem vorhin genannten Werke von der zweiten Auflage an als „Introduction" vorangestellt ist (deutsch von Tönnies 1819 und von N. Schwaiger 1886), auf die prinzipiellen Fragen ausführlich eingegangen. Unter den Franzosen war es vornehmlich Cournot, welcher in seiner „Exposition de la théorie des chances et des probabilités" 1843 (deutsch von H. Schnuse, 1849) sich diesen Fragen zuwandte und zur tieferen Begründung der Wahrscheinlichkeitstheorie beitrug. Fast zur selben Zeit (1842) ließ der

1) R. Lämmel, Untersuchungen über die Ermittlung von Wahrscheinlichkeiten (Inaug.-Dissert. 1904), p. 44—46 übt Kritik an dieser Auffassung und sieht mathematisch 1 und 0 als „Grenzwerte der Wahrscheinlichkeiten" an; Sicherheit ist für ihn nur maximale Wahrscheinlichkeit (eine recht unklare Definition). Ich kann mich seiner Anschauung nicht anschließen und möchte darauf hinweisen, daß es zum Wesen des Grenzwertes gehört, daß er nicht erreicht wird.

2) Ars conjectandi, p. 211 (Ostwalds Klassiker Nr. 108, p. 73).

3) Reflexions sur le calcul des probab. Opusc. mathém. II, 1761.

4) Essai d'Arithmétique morale. Suppl. Hist. Natur. IV, 1787.

5) Encycl. Metropol. II, Artikel „Theory of probability", 1845, p. 396.

deutsche Philosoph J. J. Fries seinen „Versuch einer Kritik der Prinzipien der Wahrscheinlichkeitsrechnung" erscheinen, worin er sich gleich Cournot gegen die bis dahin geübte rein formale oder logische Behandlung der Wahrscheinlichkeitsrechnung wendet. Ein neueres französisches Werk, J. Bertrands „Calcul des Probabilités", 1889, widmet der philosophisch-kritischen Seite des Gegenstandes eine umfangreiche Einleitung, die sich namentlich mit einzelnen Anwendungsgebieten befaßt und in ihren Schlußfolgerungen zu einer wohl allzuweit gehenden Einschränkung des Geltungsbereiches führen würde. Das Verhältnis der Wahrscheinlichkeitsrechnung zu dem hypothetisch-disjunktiven Urteil hat Ch. Sigwart in seiner „Logik" (II. Bd., 1878, 1893) klargelegt und dadurch ihre Stellung im System der Logik gekennzeichnet.

Aus der jüngsten Zeit stammen drei bemerkenswerte Arbeiten, die sich mit der philosophischen Kritik der Wahrscheinlichkeitsrechnung ausschließlich befassen, die mathematische Seite völlig beiseite lassend oder jedenfalls in zweite Linie stellend. Die Broschüre A. Ficks, betitelt „Philosophischer Versuch über die Wahrscheinlichkeiten", 1883, geht auf das Wesen der Bildung von Wahrscheinlichkeiten ein, ist aber zu allgemein gehalten, um eine für die praktische Ausführung geeignete Richtschnur zu geben. Am tiefsten geht auf alle hierher gehörigen Fragen J. v. Kries in seiner ausführlichen Schrift über „Die Prinzipien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung" 1886 ein; ihm handelt es sich vor allem um Feststellung der Bedingungen für eine begründete Wahrscheinlichkeitsaufstellung, ferner um eine Prüfung der bisher von der Wahrscheinlichkeitsrechnung gemachten Anwendungen daraufhin, wieweit hier jene Bedingungen vorhanden sind, und um die Betonung der objektiven Bedeutung auf richtiger Grundlage berechneter Wahrscheinlichkeiten; auf das X. Kapitel dieses Buches muß an dieser Stelle besonders hingewiesen werden, weil es die geschichtliche Darstellung der philosophischen Grundlagen der Wahrscheinlichkeitstheorie und eine Kritik der hierher gehörigen Versuche unternimmt. In teilweisem Gegensatze zu der Auffassung v. Kries, welche vermöge ihrer strengen Forderungen geeignet ist, den Anwendungsbereich der Wahrscheinlichkeitsrechnung einzuschränken, steht C. Stumpf in seiner 1892 in den Sitzungsberichten der philosophischen Klasse der bayrischen Akademie veröffentlichten Abhandlung „Über den Begriff der mathematischen Wahrscheinlichkeit", indem er sich auf das Prinzip des mangelnden Grundes stützt und die Zulässigkeit einer Wahrscheinlichkeitsaufstellung an viel weiter gefaßte Bedingungen knüpft. Goldschmidts Schrift: „Die Wahrscheinlichkeitsrechnung. Versuch einer Kritik", 1897, die sich zum Teil als eine Kritik der beiden vorgenannten Arbeiten darstellt, läßt an manchen Stellen die nötige Schärfe der Auffassung und Bestimmtheit des Ausdrucks vermissen. Sehr beachtenswerte Ausführungen zu den

Grundlagen der Wahrscheinlichkeitstheorie enthält das 1906 erschienene Buch von H. Bruns: „Wahrscheinlichkeitsrechnung und Kollektivmaßlehre", auf das wir in anderem Zusammenhange noch zu sprechen kommen werden.

§ 2. Direkte Wahrscheinlichkeitsbestimmung.

10. Bildung und Zählung der möglichen und günstigen Fälle. Die Bestimmung einer Wahrscheinlichkeit auf Grund der in Nr. 7 aufgestellten Definition setzt voraus, daß sich aus den Bedingungen des Problems eine Gesamtheit gleichberechtigter und innerhalb dieser die Gesamtheit der dem betreffenden Ereignis günstigen Fälle konstruieren lasse; mit der Zählung dieser Gesamtheiten ist das Problem gelöst. Die Bildung der Fälle ist ein kombinatorischer Prozeß, dessen richtige Durchführung eine scharfe Zergliederung der Bedingungen der Aufgabe erfordert. Hiermit hängt die Tatsache zusammen, daß die *Kombinatorik*, welche später zu einem selbständigen Zweige der reinen Mathematik sich entwickelt hat, mit der Wahrscheinlichkeitsrechnung zugleich ihren Anfang genommen und von ihr die ersten Aufgaben empfangen hat.[1]) Die Kombinatorik lehrt einerseits die systematische Bildung von Komplexionen gegebener Elemente, die bestimmten Bedingungen zu entsprechen haben, und befaßt sich andererseits mit der Zählung dieser Komplexionen.

11. Formeln der Kombinatorik. Die kombinatorischen Grundoperationen sind die Bildung von Permutationen, Kombinationen und Variationen.

Eine Reihe von n *verschiedenen* Elementen *permutieren* heißt, dieser Reihe alle möglichen Anordnungen geben; die Anzahl der Permutationen ist durch

$$1 \cdot 2 \cdot 3 \cdots n, \text{ symbolisch } n!,$$

gegeben. Sie vermindert sich, wenn sich unter den Elementen a *gleiche* Elemente befinden, auf

$$\frac{n!}{a!},$$

1) Pascal, der sich mit Fermat in den Ruhm der Begründung der Wahrscheinlichkeitsrechnung teilt, hat in seinem „Traité du triangle arithmétique" (1665) die erste bedeutendere Arbeit über Kombinatorik geliefert. Fast um dieselbe Zeit schrieb Leibniz seine „Dissertatio de arte Combinatoria" (1666) und handelte Wallis in seinem „Treatise of algebra" (1685) in einem besonderen Kapitel über den Gegenstand. Das erste umfassende Werk über Wahrscheinlichkeitsrechnung, Jacob Bernoullis posthume „Ars conjectandi", 1713, acht Jahre nach des Verfassers Tode von seinem Neffen Nikolaus I. Bernoulli herausgegeben, bringt in seinem zweiten Teile auch die erste systematische Darstellung der „Permutations- und Kombinationslehre"; von J. Bernoulli stammen auch einzelne der heute in der Kombinatorik üblichen Benennungen.

und sie vermindert sich weiter auf

$$\frac{n!}{a!\,b!},$$

wenn außer den a gleichen Elementen einer Art noch b gleiche Elemente einer andern Art sich vorfinden usw.

Eine Reihe von n Elementen zur r-ten Klasse *ohne Wiederholung kombinieren* heißt, sie auf alle möglichen Arten in Gruppen von je r Elementen zusammenstellen derart, daß ein Element in einer Gruppe nur einmal vorkommt; auf die Anordnung der Elemente innerhalb einer Komplexion kommt es dabei nicht an. Die Anzahl der Kombinationen ist

$$\frac{n(n-1)\cdots(n-r+1)}{1\cdot 2\cdots r}, \text{ symbolisch } \binom{n}{r},$$

und kann durch Fakultäten ausgedrückt werden, indem auch

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

ist. Die Klassenzahl ist durch die Elementenzahl beschränkt, indem $r \leqq n$ sein muß.

Eine Reihe von n Elementen zur r-ten Klasse *mit Wiederholung kombinieren* heißt, auf alle möglichen Arten Gruppen von r Elementen bilden mit der Maßgabe, daß ein und dasselbe Element in einer Komplexion mehrmal und selbst r-mal sich wiederholen kann; auf die Anordnung der Elemente innerhalb einer Komplexion kommt es nicht an. Die Anzahl solcher Kombinationen ist

$$\frac{n(n+1)\cdots(n+r-1)}{1\cdot 2\cdots r},$$

und durch Fakultäten ausgedrückt:

$$\frac{(n+r-1)!}{r!\,(n-1)!}.$$

Die Klassenzahl ist hier unbeschränkt.

Aus einer Reihe von n Elementen Variationen *ohne Wiederholung* zur r-ten Klasse bilden heißt, auf alle Arten und in allen möglichen Anordnungen Gruppen von je r verschiedenen Elementen bilden. Die Anzahl derartiger Komplexionen ist

$$n(n-1)\cdots(n-r+1) = r!\binom{n}{r},$$

durch Fakultäten ausgedrückt:

$$\frac{n!}{(n-r)!}.$$

Die Klassenbildung ist durch die Elementenanzahl beschränkt. Die Variationen der obersten, n-ten, Klasse fallen mit den Permutationen der Elemente zusammen.

Aus einer Reihe von n Elementen *Variationen mit Wiederholung* der r-ten Klasse bilden heißt, auf alle möglichen Arten und in allen Anordnungen Gruppen von je r Elementen bilden, worunter sich gleiche Elemente selbst bis zur Anzahl r befinden können. Die Anzahl dieser Komplexionen ist

$$n^r,$$

und die Klassenbildung ist unbeschränkt.

Mit einer von Kramp[1]) stammenden Bezeichnung für Produkte aufeinander folgender natürlicher Zahlen, wonach für das Produkt $n(n-1)\cdots(n-r+1)$ das Symbol $n^{r/-1}$ oder auch das Symbol $(n-r+1)^{r/1}$ gebraucht wird, läßt sich

$$n! \text{ auch in der Form } 1^{n/1}$$

$$\frac{n!}{a!\,b!} \quad \text{in der Form} \quad \frac{1^{n/1}}{1^{a/1}\,1^{b/1}}$$

$$\binom{n}{r} \quad \text{in der Form} \quad \frac{n^{r/-1}}{1^{r/1}}$$

usw. schreiben.

Bemerkenswert und wichtig ist der Umstand, daß sich die Anzahlen der verschiedenen Arten von Komplexionen, bis auf die Variationen mit Wiederholung, durch Fakultäten darstellen lassen. Demnach würde man für praktische Zwecke mit einer Tafel der Fakultäten das Auskommen finden.[2]) Indessen wachsen die Fakultäten mit der Zahl n sehr rasch an und werden bei sehr großem n physisch unberechenbar; um so weniger wären Rechnungen mit mehreren Fakultäten großer Zahlen direkt zu bewerkstelligen. Während die Fakultäten der zehn ersten Zahlen noch innerhalb mäßiger Grenzen sich bewegen — es ist

$$1! = 1, \quad 2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120,$$
$$6! = 720, \quad 7! = 5040, \quad 8! = 40320, \quad 9! = 362880! \quad 10! = 3{,}628800$$

—, ist die Faktorielle von 20 eine neunzehnziffrige Zahl:

$$20! = 2_{,,,}432902_{,,}008176{,}640000$$

und die Fakultät von 30 schon dreiunddreißigziffrig:

$$30! = 265_{,,,,,}252859_{,,,,}810729_{,,,}458636_{,,}308480{,}000000.$$

---

1) Chr. Kramp, Élements d'Arithm., 1808.

2) Unter dem Titel „Tabularum ad faciliorem probabilitatis computationem utilem Enneas" hat (1824) C. F. Degen die zwölfstelligen Logarithmen aller $n!$ von $n = 1$ bis $n = 1200$ herausgegeben. Eine sechsstellige Tafel gleichen Umfanges ist A. de Morgans Artikel „Theory of Probabilities" in der Encycl. Metropol. vol. II (1845) beigefügt.

12. Die Formel von Stirling. Für die mathematische Ausbildung der Wahrscheinlichkeitstheorie bedeutete es einen großen Fortschritt, als es gelang, eine Funktion zu finden, welche den Wert der Fakultät einer großen Zahl angenähert gibt und rechnerisch leicht zu handhaben ist. De Moivre[1]) hat den Weg hierzu angegeben und war zu einem Resultate gelangt, in welchem ihm nur noch die Natur einer in Form einer unendlichen Reihe dargestellten Konstanten unbekannt blieb; dieser vollendende Schritt gelang Stirling[2]), nach welchem später die betreffende Näherungsformel benannt worden ist.

Zum Zwecke ihrer Ableitung gehen wir von der Tatsache aus, daß sich $n!$ durch ein Eulersches Integral zweiter Gattung, durch eine Gammafunktion darstellen läßt. Es ist nämlich

$$n! = \int_0^\infty x^n e^{-x} dx. \tag{1}$$

Gelingt es, für das Integral einen angenäherten Wert zu finden, so ist dadurch das angestrebte Ziel erreicht. Dazu ist vor allem notwendig, die Natur der Funktion

$$f(x) = x^n e^{-x},$$

welche unter dem Integralzeichen steht, näher kennen zu lernen. Diese innerhalb der Integrationsgrenzen beständig positive Funktion hat für $x = 0$ den Wert 0, nähert sich mit beständig wachsendem x der Grenze Null und erlangt an der Stelle, welche sich aus der Gleichung:

$$f'(x) = x^{n-1} e^{-x}(n - x) = 0$$

ergibt, das ist für $x = n$, ihr Maximum:

$$f(n) = n^n e^{-n}. \tag{2}$$

Von diesem Maximum fällt sie sehr rasch ab und erlangt z. B. an der Stelle $x = 2n$ den Wert

$$f(2n) = (2n)^n e^{-2n},$$

der ein um so kleinerer Bruchteil ihres größten Wertes ist, je beträchtlicher n. So ist:

für $n =$ 10	$f(n) =$	453999,3	$f(2n) =$	$0{,}046\,489 \cdots f(n)$
„ „ 20	„ =	einer 18-ziffr. Zahl $216127_{,,} \cdots$	„ =	$0{,}002\,161 \cdots f(n)$
„ „ 30	„ =	einer 32-ziffr. Zahl $19_{,,,,,}2665 \cdots$	„ =	$0{,}000\,10 \cdots f(n)$
„ „ 100	„ =	einer 157-ziffr. Zahl $3_{(26)}\,72007 \cdots$	„ =	$0{,}000\,000\,000\,000\,047 \cdots f(n)$.

1) A. de Moivre, The Doctrine of Chances (1718, 1738, 1756). — Miscellanea Analytica (1730).

2) J. Stirling, Methodus differentialis (1730).

Der typische Verlauf der Funktion ist aus der Fig. 1 ersichtlich, in welcher jedoch die Ordinaten gegenüber den Abszissen in 10000-facher Verkürzung dargestellt sind; sie entspricht dem Fall $n = 10$.

Dies vorausgeschickt, führen wir in dem Integral (1) eine neue Variable t mit Hilfe der Substitution

$$x^n e^{-x} = f(n) e^{-t^2} \qquad (3)$$

ein; x wird das Intervall $(0, \infty)$ durchlaufen, während t beständig wachsend von $-\infty$ bis ∞ fortschreitet, so daß zunächst gelten wird:

$$n! = f(n) \int_{-\infty}^{\infty} e^{-t^2} \frac{dx}{dt} dt. \qquad (4)$$

Um den Differentialquotienten $\frac{dx}{dt}$ zu finden, werde eine neue Hilfsvariable ξ mittels der Gleichung

$$x = n + \xi$$

eingeführt, welche beständig wachsend das Intervall $(-n, \infty)$ durchläuft, während x auf seinem Gebiete $(0, \infty)$ stetig sich ändert; zugleich ist

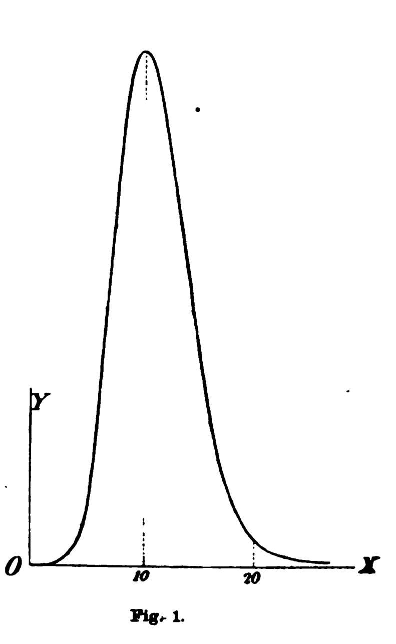


Fig. 1.

$$\frac{dx}{dt} = \frac{d\xi}{dt}. \qquad (5)$$

Aus der Gleichung (3) wird dadurch mit Rücksicht auf (2)

$$(n + \xi)^n e^{-n-\xi} = n^n e^{-n} e^{-t^2},$$

woraus

$$e^{-t^2} = \left(1 + \frac{\xi}{n}\right)^n e^{-\xi}$$

und weiter

$$t^2 = \xi - n\,\mathrm{l}\left(1 + \frac{\xi}{n}\right)$$

folgt. Entwickelt man den Logarithmus in eine Reihe, so ergibt sich für t^2 die Darstellung:

$$t^2 = \frac{\xi^2}{2n} - \frac{\xi^3}{3n^2} + \frac{\xi^4}{4n^3} - \dots \qquad (6)$$

Wählt man umgekehrt ξ in der Form

$$\xi = a_1 t + a_2 t^2 + a_3 t^3 + \dots$$

— wobei ein von t freies Glied fortgelassen wurde, weil ξ mit t zugleich verschwindet —, so ergeben sich durch Einsetzung dieses Ausdrucks bei Anwendung des Satzes der unbestimmten Koeffizienten zur Bestimmung von $a_1, a_2, \ldots$ die Gleichungen:

$$\left.\begin{array}{r} \frac{a_1^2}{2n} = 1 \\ \frac{a_1 a_2}{n} - \frac{a_1^3}{3n^2} = 0 \\ \frac{2a_1 a_3 + a_2^2}{2n} - \frac{a_1^2 a_2}{n^2} + \frac{a_1^4}{4n^3} = 0 \\ \ldots\ldots\ldots\ldots \end{array}\right\} \text{ woraus } \left\{\begin{array}{l} a_1 = \sqrt{2n} \\ a_2 = \frac{2}{3} \\ a_3 = \frac{\sqrt{2n}}{18n} \\ \ldots\ldots; \end{array}\right.$$

daß für a_1 die positive Wurzel zu nehmen ist, folgt aus der Bemerkung, daß $\left(\frac{d\xi}{dt}\right)_{t=0} = a_1$ positiv sein muß, da ξ mit t beständig wächst.

Die Gleichung (4) geht nun über in

$$n! = f(n)\int_{-\infty}^{\infty} e^{-t^2}(a_1 + 2a_2 t + 3a_3 t^2 + \ldots)\,dt,$$

und da

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-t^2}\,dt &= \sqrt{\pi} \\ \int_{-\infty}^{\infty} te^{-t^2}\,dt &= 0 \\ \int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt &= \frac{1}{2}\sqrt{\pi} \\ \ldots\ldots\ldots&, \end{aligned}$$

so erhält man nach Einsetzung dieser und der Werte von $a_1, a_3, \ldots$

$$n! = n^n e^{-n}\sqrt{2\pi n}\left(1 + \frac{1}{12n} + \ldots\right). \tag{7}$$

Dies ist die auf Glieder von der Ordnung $\frac{1}{n}$ abgekürzte *Stirlingsche Formel.* Auf einen Punkt in ihrer Ableitung muß hingewiesen werden. Die Entwicklung (6) ist nur so lange konvergent, als $\left|\frac{\xi}{n}\right| < 1$, also ξ in dem Intervall $(-n, n)$, x in dem Intervall $(0, 2n)$ und t in dem zugeordneten Intervall $(-\infty, t_{2n})$ verbleibt; die später vollzogene Integration erstreckt sich aber auf das ganze Gebiet $(0, \infty)$ von x, respektive $(-\infty, \infty)$ von t; dies hat zur Folge, daß die

Reihe in (7) divergent wird. Nichtsdestoweniger ist sie zur näherungsweisen Berechnung von $n!$ geeignet; der Grund liegt in der eingangs hervorgehobenen Tatsache, daß die Funktion $f(x)$ in dem Intervall $(0, 2n)$ Werte annimmt, welche alle übrigen in um so höherem Maße überschreiten, je größer n ist.

Für die Zwecke der Wahrscheinlichkeitsrechnung genügt es fast durchwegs, die Stirlingsche Formel in der weitergehenden Abkürzung

$$n! = n^n e^{-n} \sqrt{2\pi n} \tag{8}$$

zu verwenden, in welcher die rechte Seite immer einen zu kleinen Wert für die links stehende Fakultät liefert.

Um die Natur dieser merkwürdigen Näherungsformel erkennen zu lassen und zu zeigen, daß sie schon bei mäßigen Werten von n gut verwendbar ist, führen wir die folgenden Beispiele an: Es ist

$$10! = 3\,628\,800$$
$$10^{10} e^{-10} \sqrt{20\pi} = 3\,598\,699$$

Differenz = 30 101 = 0,008 ... des richtigen Wertes;

$$20! = 2_{'''}432902_{''}008176{,}640000$$
$$20^{20} e^{-20} \sqrt{40\pi} = 2_{'''}422786_{''}385510{,}400000$$

die Differenz = $10115_{''}622666{,}240000$

macht 0,004 ... des richtigen Wertes aus;

$$30! = 265_{''''}252859_{'''}\ldots$$
$$30^{30} e^{-30} \sqrt{60\pi} = 264_{''''}517093_{'''}\ldots$$

Differenz = $735766_{'''}\ldots$ = 0,0028 . . des richtigen Wertes.

Der absolute Fehler, den man bei Anwendung der Formel (8) begeht, wächst also mit n ins Ungeheure, der relative Fehler aber nimmt mit wachsendem n ab und ist selbst bei $n = 10$ so klein, daß er für praktische Zwecke fast ausnahmslos außer Betracht bleiben kann.

13. Beispiel I. *Die Wahrscheinlichkeit zu bestimmen, daß bei zweimaligem Aufwerfen einer Münze Wappen erscheint.*

Bezeichnet man die Wappenseite mit W, die Schriftseite mit S, so kann das Ergebnis zweier Würfe einer der Fälle sein:

$$WW; \quad WS; \quad SW; \quad SS;$$

da drei dieser Fälle dem erwarteten Ereignis günstig sind, so ist dessen Wahrscheinlichkeit

$$\mathfrak{W}(W) = \frac{3}{4}.$$

In anderer — unrichtiger — Weise faßte D'Alembert[1]) die Aufgabe auf; mit dem Hinweise, daß, sofern Wappen im ersten Wurf fällt, das Spiel beendet und der zweite Wurf unnötig sei, setzt er die Fälle

$$W;\quad SW;\quad SS$$

und rechnet die Wahrscheinlichkeit mit $\frac{2}{3}$; daß der erste Fall nicht gleichwertig den beiden andern ist, gab D'Alembert nicht zu.

14. Beispiel II. *Es ist die Wahrscheinlichkeit zu ermitteln, daß man mit zwei Würfeln eine bestimmte Summe treffen werde.*

Die möglichen Fälle werden durch die $6^2 = 36$ Variationen mit Wiederholung gebildet, welches die sechs Elemente 1, 2, ... 6 zulassen.

Von diesen führt je eine (1, 1; 6, 6) die Summen 2 und 12 herbei; je zwei (1, 2; 2, 1; — 5, 6; 6, 5) ergeben die Summen 3 und 11; je drei (1, 3; 2, 2; 3, 1; — 4, 6; 5, 5; 6, 4) die Summen 4 und 10; je vier die Summen 5 und 9; je fünf die Summen 6 und 8 und sechs (1, 6; 2, 5; 3, 4; 4, 3; 5, 2; 6, 1) liefern die Summe 7. Wird also mit $W(s)$ die Wahrscheinlichkeit der Summe s bezeichnet, so ist

$$\mathfrak{W}(2) = \mathfrak{W}(12) = \frac{1}{36};\quad \mathfrak{W}(3) = \mathfrak{W}(11) = \frac{1}{18};\quad \mathfrak{W}(4) = \mathfrak{W}(10) = \frac{1}{12};$$

$$\mathfrak{W}(5) = \mathfrak{W}(9) = \frac{1}{9};\quad \mathfrak{W}(6) = \mathfrak{W}(8) = \frac{5}{36};\quad \mathfrak{W}(7) = \frac{1}{6}.$$

Wenn Leibniz[2]) der Summe 11 nur einen und der Summe 7 drei Fälle zuschrieb, so schwebten ihm die Kombinationen mit Wiederholung als Vertreter der gleichmöglichen Fälle vor, und diese Auffassung ist irrig; denn während z. B. der Fall 6, 6 nur auf eine Art zustande kommen kann, sind für 5, 6 zwei Entstehungsweisen möglich, indem sich die beiden Würfel in die Nummern 5, 6 auf zwei Arten verteilen können.

15. Beispiel III. *Die Wahrscheinlichkeit zu bestimmen, daß mit drei Würfeln die Summe 9, beziehungsweise 10, 11, 12 fallen werde.*

Wie im vorigen Beispiel sind die möglichen Fälle durch die Variationen mit Wiederholung der Elemente 1, 2, ..., 6 dargestellt, deren Anzahl, da es sich nun um drei Würfel handelt, $6^3 = 216$ ist.

Die Entstehungsarten der bezeichneten Summen sind die folgenden:

---

1) Artikel „Croix ou Pile" in der „Encyclopédie" (1754).

2) Dissertatio de Arte Combinatoria (1666).

Summe 9:	Fälle:	Summe 10:	Fälle:	Summe 11:	Fälle:	Summe 12:	Fälle:
126	6	136	6	146	6	156	6
135	6	145	6	155	3	246	6
144	3	226	3	236	6	255	3
225	3	235	6	245	6	336	3
234	6	244	3	335	3	345	6
333	1	334	3	344	3	444	1
	25		27		27		25

Wiewohl sich für jede dieser Summen, wenn man bloß auf die Größe der Summanden achtet, sechs Entstehungsarten ergeben, so repräsentieren diese doch nicht auch gleichmögliche Einzelfälle; so kann die Verbindung 333 aus der ersten Kolonne nur auf eine Art, die Verbindung 225 hingegen auf drei Arten, die Verbindung 234 aber auf sechs Arten entstehen, indem sich die Nummern 2, 2, 5, respektive 2, 3, 4 auf drei, beziehungsweise sechs Arten auf die drei Würfel verteilen können. Mithin ist, wenn man mit p_s die Wahrscheinlichkeit der Summe s bezeichnet,

$$p_9 = p_{12} = \frac{25}{216}, \quad p_{10} = p_{11} = \frac{27}{216} = \frac{1}{8}.$$

An diese Aufgabe knüpft sich die erste Untersuchung über eine wahrscheinlichkeitstheoretische Frage, die die Literatur nachweist. Diese Frage bezog sich auf das sogenannte Knöchelspiel (passe-dix), bei welchem es sich darum handelt, mit drei Würfeln eine Summe über zehn zu werfen. Ein fleißiger Beobachter dieses Spieles hatte die Wahrnehmung gemacht, daß die Summe 11 häufiger erscheint als 12 und 10 öfter als 9; da nach seiner Auffassung alle vier Summen auf gleich viele Arten entstehen konnten, verlangte er von Galilei[1]) Aufklärung über diesen Widerspruch zwischen Beobachtung und Rechnung. Galilei gab sie ihm in der oben dargelegten Weise.

16. Beispiel IV. *Es liegen drei äußerlich gleiche Kästchen A, B, C mit je zwei Lädchen vor; im Kästchen A enthält jedes Lädchen eine Goldmünze, in B jedes eine Silbermünze, in C das eine eine Gold-, das andere eine Silbermünze. Man greift ein Kästchen beliebig heraus; welches ist die Wahrscheinlichkeit, daß es C sei? — Man öffnet ein Lädchen und besichtigt den Inhalt; wie groß ist nach dieser Kenntnisnahme die Wahrscheinlichkeit, daß man das Kästchen C erwählt habe?*

Die Antwort auf die erste Frage ist leicht gegeben; von drei gleichmöglichen Fällen ist einer dem bezeichneten Ereignis günstig, die Wahrscheinlichkeit ist $p = \frac{1}{3}$.

---

1) G. Galilei, Considerazione sopra il giuoco dei dadi (unbekannten Datums, jedenfalls vor 1642); abgedruckt in „Le Opere di Galileo Galilei", vol. XIV, Firenze 1855.

Bezüglich der zweiten Frage liegt ein Fehlschluß nahe. Was auch das geöffnete Lädchen enthalten möge, es kommen immer nur noch zwei Kästchen in Frage, weil eines durch das erlangte Wissen ausgeschlossen ist, und diese Überlegung könnte dazu führen, die Wahrscheinlichkeit, daß man das Kästchen C herausgegriffen habe, mit $\frac{1}{2}$ anzusetzen.

Indessen muß hier anders geschlossen werden. Enthält das geöffnete Lädchen eine Goldmünze, so kann man es allerdings nur mit dem Kästchen A oder mit C zu tun haben; diese beiden Kästchen enthalten aber drei Lädchen mit Goldmünzen, A deren zwei, C nur eines; die Frage nach der Wahrscheinlichkeit, daß man das Kästchen C vor sich habe, ist also gleichbedeutend mit der Frage, daß man von den drei Lädchen mit Goldmünzen dasjenige geöffnet habe, welches sich in C befindet, und diese Wahrscheinlichkeit ist, weil unter drei Fällen ein günstiger vorkommt, $\frac{1}{3}$, dieselbe wie im ersten Falle.[1]) Zu dem gleichen Resultat wird man geführt, wenn das geöffnete Lädchen eine Silbermünze enthält.

Es liegt nur in der eigentümlichen Konstruktion des Beispiels, daß das durch das Öffnen eines Lädchens erworbene, gegenüber der ersten Fragestellung vermehrte Wissen an der Wahrscheinlichkeit des erwarteten Ereignisses nichts ändert. Enthielten die Kästchen je drei Lädchen, diese drei Lädchen in A Gold-, in B Silbermünzen, während in C zwei Gold- und eine Silbermünze auf die Lädchen verteilt wären, so verhielte es sich anders. Wieder wäre $\frac{1}{3}$ die Wahrscheinlichkeit dafür, daß man bei vollzogener Wahl das Kästchen C ergreift; findet man aber in einem geöffneten Lädchen eine Goldmünze, so gehört das Lädchen mit der Wahrscheinlichkeit $\frac{2}{5}$ dem Kästchen C an, und mit der Wahrscheinlichkeit $\frac{1}{4}$, wenn die geöffnete Lade eine Silbermünze enthielt.

17. Beispiel V. *Zwei Personen, A und B, gleich geschickt, werfen Kugeln nach einem Ziel; derjenige gewinnt, welcher dem Ziele am nächsten kommt. A hat zwei, B eine Kugel zu werfen. Welches ist die Wahrscheinlichkeit, daß A gewinnen werde?*

Bezeichnet man die Kugeln des A mit a_1, a_2, die Kugel des B mit b, so können folgende Permutationen der Kugel mit Rücksicht auf ihre Entfernung vom Ziel eintreten:

---

1) Vgl. hiermit die Schlußweise bei J. Bertrand, Calcul des Probabilités (1889) p. 2, und bei H. Poincaré, Calcul des Probabilités (1896) p. 3. — Ich benütze diese Gelegenheit, um meinen eigenen Fehlschluß in dem Bericht über „Die Entwicklung der Wahrscheinlichkeitstheorie und ihrer Anwendungen" (1899), Jahresber. d. Deutschen Mathem.-Verein. VII. zu berichtigen.

$$a_1 a_2 b, \quad a_1 b a_2, \quad a_2 a_1 b, \quad a_2 b a_1, \quad b a_1 a_2, \quad b a_2 a_1;$$

ist jedesmal die letzte Kugel die dem Ziel nächste, so sind unter diesen sechs gleichmöglichen Fällen vier der Person A günstig; für sie beträgt daher die Wahrscheinlichkeit, zu gewinnen, $\frac{2}{3}$.

Man kann auch folgenden Weg zur Lösung einschlagen. Jede der Kugeln a_1, a_2 kann näher oder weiter fallen als die Kugel b; dies gibt folgende Verbindungen:

1) a_1 weiter als b, a_2 weiter als b;
2) a_1 „ „ b, a_2 näher „ b;
3) a_1 näher „ b, a_2 weiter „ b;
4) a_1 „ „ b, a_2 näher „ b.

Die Verbindungen 2), 3), 4) sind dem A günstig; aber der Schluß, daß seine Wahrscheinlichkeit $\frac{3}{4}$ betrage, wäre falsch. Denn diese Verbindungen stellen nicht gleichmögliche Fälle dar; vielmehr zerfallen 1) und 4) in je zwei Fälle, da im Falle 1) die Reihenfolge der Kugeln $a_1 a_2 b$ oder $a_2 a_1 b$, im Falle 4) $b a_1 a_2$ oder $b a_2 a_1$ sein kann. Wird dies beachtet, so ergibt sich wieder $p = \frac{2}{3}$.

13. Beispiel VI. *Die Wahrscheinlichkeit zu bestimmen, daß eine Münze in n Würfen abwechselnd Wappen und Schrift zeigt.*

Die möglichen Fälle sind die 2^n Varationen mit Wiederholung der Elemente W, S.

Unter diesen Variationen gibt es nur zwei, nämlich $WSWS\cdots$ und $SWSW\cdots$, welche dem erwarteten Ereignis günstig sind; seine Wahrscheinlichkeit ist daher $p = \frac{2}{2^n} = \frac{1}{2^{n-1}}$.

Wäre festgesetzt worden, daß Wappen an erster Stelle erscheinen solle, so gäbe es nur eine günstige Verbindung, die Wahrscheinlichkeit wäre die Hälfte der vorigen.

Überhaupt ist $\frac{1}{2^n}$ die Wahrscheinlichkeit für *irgend* eine *bestimmte* Anordnung von Wappen und Schrift. Es hat also die durch ihre Gesetzmäßigkeit auffallende Anordnung $WSWS\cdots$ keine geringere Wahrscheinlichkeit als irgend eine andere unregelmäßige, aber im voraus festgesetzte Reihenfolge der Münzseiten. Wenn hiernach das Eintreffen von $WSWS\cdots$ als ein *außergewöhnliches* Ereignis bezeichnet wird, so ist diese Bezeichnung insofern unzutreffend, als sie mit dem gleichen Rechte *jeder* andern Anordnung gebührte, wenn diese im voraus bezeichnet worden wäre. Anders verhält es sich, wenn man das angeführte regelmäßige Ereignis gegenüberhält der übergroßen

Anzahl derjenigen, welche keine in die Augen springende Regelmäßigkeit aufweisen. Wenn jemand aus einer Urne, die auf neun gleichen Kärtchen die Buchstaben *b, e, i, l, l, n, o, r, u* enthält, diese Kärtchen sukzessive herausziehend sie in der Ordnung *bernoulli* erhielte, so würde er dies mit Rücksicht darauf, daß die Buchstaben in dieser Aufeinanderfolge einen ihm geläufigen Namen darstellen, als ein ungewöhnliches Ereignis bezeichnen; einem andern, dem dieser Name völlig fremd ist, wird nur mehr die Lesbarkeit der Permutation auffallen, und sie wird ihm von diesem Gesichtspunkte nicht ungewöhnlicher erscheinen als etwa *rubellion* oder *bulinerol*; an sich aber stehen diese Anordnungen auf ganz gleicher Stufe mit jeder andern, wie etwa *lnirolbeu*: jede ist unter den 181 440 verschiedenen Permutationen einzig in ihrer Art und hat die Wahrscheinlichkeit $\frac{1}{181\,440}$.

19. Beispiel VII. *Eine Urne enthält a weiße und b schwarze Kugeln; man nimmt r* ($\leqq a$) *Kugeln auf einmal oder nacheinander, im letzteren Falle, ohne die gezogenen zurückzulegen, heraus; wie groß ist die Wahrscheinlichkeit, daß sie weiß seien?*

Die möglichen Fälle werden durch die Kombinationen ohne Wiederholung der $a+b$ Kugeln zur r-ten Klasse gebildet, ihre Anzahl ist also $\binom{a+b}{r} = \frac{(a+b)!}{r!\,(a+b-r)!}$.

Günstige Fälle stellen die gleichartigen Kombinationen der a weißen Kugeln vor, deren Anzahl $\binom{a}{r} = \frac{a!}{r!\,(a-r)!}$ ist.

Die verlangte Wahrscheinlichkeit ist demnach gleich

$$\frac{a^{r/-1}}{(a+b)^{r/-1}}.$$

In dem gewöhnlichen Zahlenlotto sind 9 einziffrige und 81 zweiziffrige Nummern; fünf Nummern werden gezogen; die Wahrscheinlichkeit, daß das Ergebnis einer Ziehung aus lauter einziffrigen Nummern bestehe, ist hiernach:

$$\frac{9^{5/-1}}{90^{5/-1}} = \frac{9\cdot 8\cdot 7\cdot 6\cdot 5}{90\cdot 89\cdot 88\cdot 87\cdot 86} = 0{,}0000029,$$

und die Wahrscheinlichkeit einer Ziehung aus durchwegs zweiziffrigen Nummern:

$$\frac{81^{5/-1}}{90^{5/-1}} = \frac{81\cdot 80\cdot 79\cdot 78\cdot 77}{90\cdot 89\cdot 88\cdot 87\cdot 86} = 0{,}582981.$$

20. Beispiel VIII. *Eine Urne enthält a weiße, b schwarze und c rote Kugeln. Man macht* $\alpha+\beta+\gamma$ *Ziehungen, legt die gezogene Kugel jedesmal in die Urne zurück und vermengt die Kugeln. Es wird*

die Wahrscheinlichkeit verlangt, 1) *daß zuerst* α*-mal weiß, dann* β*-mal schwarz und endlich* γ*-mal rot gezogen werde;* 2) *daß die Farben in denselben Anzahlen und geschlossen, jedoch in was immer für einer Reihenfolge erscheinen;* 3) *daß weiße, schwarze und rote Kugeln in der bezeichneten Anzahl überhaupt erscheinen.*

1) Die möglichen Fälle sind durch die $(a+b+c)^{\alpha+\beta+\gamma}$ Variationen mit Wiederholung sämtlicher Kugeln zur Klasse $\alpha+\beta+\gamma$ dargestellt.

Aus den a^α Variationen der weißen, b^β Variationen der schwarzen und c^γ Variationen der roten Kugeln in der durch die Anzahl bezeichneten Klasse entstehen $a^\alpha b^\beta c^\gamma$ dem erstgedachten Ereignis günstige Verbindungen; seine Wahrscheinlichkeit ist daher

$$p_1 = \frac{a^\alpha b^\beta c^\gamma}{(a+b+c)^{\alpha+\beta+\gamma}}.$$

2) Unter der gleichen Anzahl möglicher Fälle gibt es

$$a^\alpha b^\beta c^\gamma + a^\alpha c^\gamma b^\beta + b^\beta a^\alpha c^\gamma + b^\beta c^\gamma a^\alpha + c^\gamma a^\alpha b^\beta + c^\gamma b^\beta a^\alpha = 6a^\alpha b^\beta c^\gamma$$

günstige Fälle; die sechs Produkte entsprechen den sechs möglichen Anordnungen der drei Farben; mithin ist

$$p_2 = \frac{6a^\alpha b^\beta c^\gamma}{(a+b+c)^{\alpha+\beta+\gamma}} = 6p_1.$$

3) Aus jeder der $a^\alpha b^\beta c^\gamma$ Verbindungen von α weißen, β schwarzen und γ roten Kugeln lassen sich $\frac{(\alpha+\beta+\gamma)!}{\alpha!\,\beta!\,\gamma!}$ verschiedene Permutationen bilden; folglich ist

$$\frac{(\alpha+\beta+\gamma)!}{\alpha!\,\beta!\,\gamma!}\,a^\alpha b^\beta c^\gamma$$

die Anzahl aller Anordnungen, in welchen α weiße, β schwarze und γ rote Kugeln in irgend welcher Reihenfolge erscheinen können, und somit, da die Anzahl der möglichen Fälle dieselbe ist wie in den vorhergehenden Aufgaben, ist

$$p_3 = \frac{(\alpha+\beta+\gamma)!}{\alpha!\,\beta!\,\gamma!}\,\frac{a^\alpha b^\beta c^\gamma}{(a+b+c)^{\alpha+\beta+\gamma}}.$$

Beispielsweise ist für $a=2,\ b=3,\ c=4$ und $\alpha=1,\ \beta=2,\ \gamma=3$

$$p_1 = 0{,}00217,\quad p_2 = 0{,}01300,\quad p_3 = 0{,}1302.$$

21. Beispiel IX. *Aus einer Urne, welche n gleiche Kugeln enthält, wird ein Teil oder werden alle gezogen. Es ist die Wahrscheinlichkeit zu bestimmen, daß eine gerade Anzahl von Kugeln herauskommt.*

Da *eine* Kugel auf so viele Arten gezogen werden kann, als es Kugeln gibt, *zwei* Kugeln auf so viele Arten, als die n Kugeln

Kombinationen zur zweiten Klasse ohne Wiederholung zulassen usw., so ist die Anzahl der möglichen Fälle

$$m = n + \binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n};$$

günstig sind darunter die Kombinationen zu einer geraden Anzahl, deshalb ist

$$g = \binom{n}{2} + \binom{n}{4} + \cdots$$

die Anzahl der günstigen Fälle. Nun ist aber

$$(1+1)^n = 1 + n + \binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n}$$

$$(1-1)^n = 1 - n + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^n \binom{n}{n};$$

aus der ersten Gleichung ergibt sich:

$$m = 2^n - 1 \tag{1}$$

und durch Addition beider:

$$g = 2^{n-1} - 1. \tag{2}$$

Mithin ist die Wahrscheinlichkeit, eine gerade Anzahl von Kugeln zu ziehen,

$$p = \frac{2^{n-1} - 1}{2^n - 1}; \tag{3}$$

für eine ungerade Anzahl:

$$q = 1 - p = \frac{2^{n-1}}{2^n - 1}. \tag{4}$$

Es ist also, bei jedem n, $p < q$. Dieses auf den ersten Blick befremdliche Resultat, daß die ungeraden Anzahlen den geraden gegenüber bevorzugt sind, wird sofort klar, wenn man zur unteren Grenze herabgeht; enthält die Urne nur eine Kugel, so kann nur eine ungerade Anzahl gezogen werden, es ist $p = 0$, $q = 1$; bei zwei Kugeln bestehen zwei Möglichkeiten für das Ziehen einer ungeraden Anzahl (einer Kugel) und nur eine für das Ziehen einer geraden Anzahl, es wird $p = \frac{1}{3}$, $q = \frac{2}{3}$. Dieses Überwiegen des q über das p besteht fort, und erst für $n = \infty$ würde $p = q = \frac{1}{2}$.

Oben wurde vorausgesetzt, daß die Anzahl n der Kugeln bekannt sei; wüßte man bloß, daß sie n nicht übertreffen und jeden Wert von 1 bis n (mit Einschluß dieser Grenze) gleich leicht haben könne, so hätte man mit Rücksicht auf (1) und (2)

$$m = \sum_1^n (2^x - 1) = \sum_1^n 2^x - n = 2^{n+1} - n - 2$$

$$g = \sum_1^n (2^{x-1} - 1) = 2^n - n - 1;$$

demzufolge wäre dann

$$p = \frac{2^n - n - 1}{2^{n+1} - n - 2} \tag{5}$$

$$q = \frac{2^n - 1}{2^{n+1} - n - 2}, \tag{6}$$

also wieder $p < q$.

22. Beispiel X. *Aus einer Urne mit n weißen und n schwarzen Kugeln wird irgend eine gerade Anzahl von Kugeln gezogen; wie groß ist die Wahrscheinlichkeit, daß gleich viel weiße und schwarze darunter sind?*

Es werden, den Bedingungen der Aufgabe gemäß, entweder 2 oder 4 oder $\cdots$ oder $2n$ Kugeln gezogen; das erste kann auf so viele Arten geschehen, als sich aus den $2n$ Kugeln Kombinationen ohne Wiederholung zur zweiten Klasse bilden lassen; das zweite auf so viele Arten, als es derlei Kombinationen zur vierten Klasse gibt usf.; demnach ist

$$m = \binom{2n}{2} + \binom{2n}{4} + \cdots + \binom{2n}{2n}$$

die Anzahl der möglichen Fälle.

Ein günstiger Fall entsteht, so oft eine der weißen mit einer der schwarzen Kugeln sich verbindet, was auf n^2 Arten geschehen kann; ebenso, so oft eine Verbindung von zwei weißen Kugeln mit einer Verbindung von zwei schwarzen Kugeln sich vereinigt, was auf $\binom{n}{2}^2$ Arten eintreten kann usf.; die Anzahl der günstigen Fälle ist also

$$g = n^2 + \binom{n}{2}^2 + \binom{n}{3}^2 + \cdots + \binom{n}{n}^2.$$

Der Ausdruck für m entsteht aus dem Ausdruck für g im vorigen Beispiel, wenn man hier n durch $2n$ ersetzt; demzufolge ist

$$m = 2^{2n-1} - 1.$$

Um für g einen einfacheren Ausdruck zu gewinnen, beachte man, daß bei der Multiplikation der Entwicklungen von $(1+t)^n$ und $\left(1 + \frac{1}{t}\right)^n$ die von t freien Glieder $1 + n^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2$, also $g + 1$

ergeben; diese Glieder entstehen auch dadurch, daß man $(1+t)^n\left(1+\frac{1}{t}\right)^n$ auf die Gestalt

$$\frac{(1+t)^{2n}}{t^n}$$

bringt, den Zähler entwickelt und dasjenige Glied desselben, welches mit t^n behaftet ist, durch t^n dividiert; das Resultat dieser Division ist aber der Koeffizient $\binom{2n}{n}$; folglich ist

$$g+1=\binom{2n}{n}, \quad \text{woraus } g=\frac{(2n)!}{(n!)^2}-1.$$

Die verlangte Wahrscheinlichkeit ist also

$$p=\frac{\frac{(2n)!}{(n!)^2}-1}{2^{2n-1}-1}. \tag{1}$$

Für kleine Werte von n bereitet die Ausrechnung von p keine Mühe; so ist

$$\begin{array}{llll} \text{für } n=1 & p=1 \\ \text{„} \quad \text{„} =2 & \text{„} =\frac{5}{7}=0{,}7142\cdots \\ \text{„} \quad \text{„} =3 & \text{„} =\frac{19}{31}=0{,}6129\cdots \\ \text{„} \quad \text{„} =4 & \text{„} =\frac{69}{127}=0{,}5433\cdots; \end{array}$$

für große Werte von n scheitert aber die Ausführung der Formel an den darin vorkommenden Fakultäten. In solchen Fällen kommt nun die Stirlingsche Formel zur Geltung; nimmt man für $(2n)!$ den Näherungswert $(2n)^{2n}e^{-2n}\sqrt{4n\pi}$, für $n!$ entsprechend $n^n e^{-n}\sqrt{2n\pi}$ und läßt im Zähler und Nenner die subtraktive Einheit neben der andern großen Zahl weg, so ergibt sich für p der Näherungswert

$$\frac{2}{\sqrt{\pi n}}; \tag{2}$$

selbst für $n=4$ liefert dieser $0{,}5641\cdots$ in naher Übereinstimmung mit dem strengen Werte.

Mit wachsendem n nähert sich p der Grenze 0.

23. Beispiel XI. *Die Wahrscheinlichkeit zu bestimmen, mit einem Würfel in n Würfen mindestens einmal die Seite 6 zu werfen.*

Das bezeichnete Ereignis ist der Gegensatz davon, in n Würfen immer nur eine der Seiten $1, 2, \cdots, 5$ zu treffen; wird die Wahrscheinlichkeit dieses letzteren Ereignisses mit q bezeichnet, so ist die des erstgenannten $p = 1 - q$.

Nun können die sechs Würfelseiten in n Würfen sich auf 6^n Arten miteinander verbinden — dies die möglichen Fälle —, während die Seiten 1 bis 5 sich auf 5^n Arten vereinigen können — dies die günstigen Fälle des zweiten Ereignisses —; folglich ist $q = \frac{5^n}{6^n}$ und

$$p = 1 - \frac{5^n}{6^n}. \tag{1}$$

Ist p gegeben, so läßt sich aus dieser Gleichung die Anzahl der Würfe berechnen, innerhalb deren mit eben dieser Wahrscheinlichkeit das zum mindesten einmalige Eintreffen von 6 zu erwarten ist; man findet

$$n = \frac{\log(1-p)}{\log 5 - \log 6}, \tag{2}$$

die Logarithmen aus einem beliebigen System genommen; insbesondere ist für $p = \frac{1}{2}$:

$$n = \frac{\log 2}{\log 6 - \log 5} = 3{,}8 \cdots,$$

so daß man *mit Vorteil* Eins gegen Eins wetten darf, daß in vier Würfen die Seite 6 wenigstens einmal fallen werde.

24. Beispiel XII. *Die Wahrscheinlichkeit zu bestimmen, mit zwei Würfeln in n Würfen mindestens einmal den Pasch von 6 zu werfen.*

Unter den 36 möglichen Verbindungen der Würfelseiten, welche in *einem* Wurf eintreten können, gibt es deren 35, welche das erwartete Ereignis *nicht* herbeiführen; demnach gibt es in n Würfen 36^n mögliche und 35^n ungünstige Fälle, so daß die Wahrscheinlichkeit, es werde Pasch 6 nicht erscheinen, $\frac{35^n}{36^n}$ ausmacht. Die Wahrscheinlichkeit, es werde dieses Ereignis überhaupt, also mindestens einmal sich zutragen, ist daher

$$p = 1 - \frac{35^n}{36^n}. \tag{1}$$

In

$$n = \frac{\log(1-p)}{\log 35 - \log 36} \tag{2}$$

Würfen darf man also mit der Wahrscheinlichkeit p erwarten, daß wenigstens einmal beide Würfel 6 zeigen; da für $p = \frac{1}{2}$

$$n = \frac{\log 2}{\log 36 - \log 35} = 24{,}6 \cdots,$$

so darf erst bei 25 Würfen mit Vorteil Eins gegen Eins darauf gewettet werden; bei 24 Würfen wäre eine solche Wette unvorteilhaft.

Die beiden letzten Aufgaben dürfen ein historisches Interesse beanspruchen, weil sie zu den ältesten gehören, die zu mathematischen Untersuchungen über Fragen der Wahrscheinlichkeit Anlaß gaben. Ein Freund der Glücksspiele, namens Chevalier de Meré, der, obwohl kein Mathematiker von Fach, es doch verstand, anregende Fragen auf diesem Gebiete aufzuwerfen, sprach (1654) Pascal gegenüber sein Erstaunen darüber aus, daß man bei einem Würfel auf 6 in vier Würfen mit Vorteil, bei zwei Würfeln aber auf Pasch 6 in 24 Würfen nicht mit Vorteil wetten könne, vermeinend, die Zahl der nötigen Würfe sollte der Anzahl der bei einem Wurfe möglichen Fälle proportional sein; da nun bei einem Würfel 6, bei zwei Würfeln 36 Fälle stattfinden, so ist es die Zahl 24, welche sich in die Proportion $4 : 6 = 24 : 36$ einfügt. Die obige Rechnung zeigt die Unstichhaltigkeit dieser Schlußfolge.

25. Beispiel XIII. *Zwischen zwei Personen A, B soll durch die Wahl entschieden werden; für A seien a, für B seien b Stimmzettel abgegeben und in einer Urne vereinigt worden. Angenommen, es sei $a > b$, so daß A schließlich als der Gewählte erscheinen wird; wie groß ist die Wahrscheinlichkeit, daß er während des Skrutiniums ständig in der Majorität bleibe?*

Die möglichen Fälle werden durch die verschiedenen Permutationen gezählt, in welchen die Zettel aus der Urne hervorgehen können; ihre Anzahl ist also

$$m = \frac{(a+b)!}{a!\,b!}.$$

Ungünstig sind alle Permutationen, in welchen es einmal zur Stimmengleichheit kommt.

Dahin gehören zunächst alle Permutationen, welche mit einem auf B lautenden Zettel beginnen; ihre Anzahl ergibt sich, wenn man nach Wegnahme eines Zettels mit dem Namen B alle übrigen permutiert, ist also

$$\frac{(a+b-1)!}{a!\,(b-1)!}.$$

Aus jeder Permutation, welche mit A beginnt und einmal zur Stimmengleichheit führt, läßt sich auf eindeutige Weise eine mit B anhebende Permutation, aber auch umgekehrt aus jeder Permutation der letzteren Art eine solche gewinnen, die mit A anfängt und zur Stimmengleichheit führt.

Es sei z. B. $a = 5$, $b = 3$; die Permutation

$$AABBAABA$$

beginnt mit A und erreicht nach vier Zetteln Stimmengleichheit; löst man diese vier Stimmen $AABB$, die notwendig mit B aufhören, ab, setzt sie mit Weglassung dieses letzten B an das Ende, während man das ausgeschiedene B an den Anfang stellt, so erhält man in

$$BAABAAAB$$

eine mit B beginnende Permutation. Geht man hingegen von einer mit B beginnenden Permutation, wie etwa

$$BABAAAAB$$

aus, zählt von rückwärts so weit, bis A eine Stimme mehr als B hat, was notwendig einmal eintreten und mit einer auf A lautenden Stimme schließen muß, löst die Gruppe — hier AAB — ab und setzt sie an den Anfang, so ergibt sich eine mit A anfangende Permutation, in der einmal Stimmengleichheit eintritt, hier die Permutation

$$AABBABAA.$$

Mithin bestehen die ungünstigen Fälle aus zwei Gruppen je vom Umfange

$$\frac{(a+b-1)!}{a!\,(b-1)!};$$

die der gesuchten entgegengesetzte Wahrscheinlichkeit ist daher

$$q = 2\,\frac{(a+b-1)!}{a!\,(b-1)!} : \frac{(a+b)!}{a!\,b!} = \frac{2b}{a+b},$$

mithin die gesuchte Wahrscheinlichkeit:

$$p = 1 - q = \frac{a-b}{a+b};$$

je kleiner also die Majorität im Vergleich zur gesamten Stimmenzahl, desto geringer die Wahrscheinlichkeit, daß sie während des Skrutiniums anhalte.[1])

26. Beispiel XIV. *Eine Urne enthält n mit den Nummern $1, 2 \cdots n$ bezeichnete Kugeln. Man zieht diese nach und nach einzeln aus der Urne, ohne sie zurückzulegen, und verlangt die Wahrscheinlichkeit, daß keine von r ($\leq n$) ins Auge gefaßten Kugeln in der ihrer Nummer entsprechenden Ziehung erscheine.*

---

1) Vgl. hiermit die Darstellung bei D. André, Compt. rend. 105 (1887) p. 436, und bei J. Bertrand, l. c., p. 18.

Die Anzahl der möglichen Fälle ist

$$m = n!$$

entsprechend den Permutationen, in welchen die Kugeln aus der Urne hervorgehen können.

Unter den Permutationen von n Elementen gibt es aber $(n-1)!$ solche, in welchen das erste bezeichnete Element i an seiner richtigen Stelle steht; verbleiben daher

$$\varphi_n(1) = n! - (n-1)!$$

Permutationen, wo i nicht an seinem Platze erscheint.

Aus diesen sind diejenigen auszuscheiden, in welchen das zweite bezeichnete Element k an seinem Platze steht; hält man es dort fest, so bilden die $n-1$ übrigen Elemente

$$\varphi_{n-1}(1) = (n-1)! - (n-2)!$$

Permutationen, in welchen i nicht an seinem Platze ist; es verbleiben also

$$\varphi_n(2) = \varphi_n(1) - \varphi_{n-1}(1) = n! - 2(n-1)! + (n-2)!$$

Permutationen, in welchen weder i noch k an seinem Platze ist.

Aus diesen müssen diejenigen Permutationen ausgeschieden werden, in welchen das dritte bezeichnete Element l an der seinem Range entsprechenden Stelle steht; hält man es dort fest, so bilden die $n-1$ übrigen Elemente

$$\varphi_{n-1}(2) = (n-1)! - 2(n-2)! + (n-3)!$$

Permutationen, in welchen weder i noch k seinen Platz einnimmt; es verbleiben also

$$\varphi_n(3) = \varphi_n(2) - \varphi_{n-1}(2) = n! - 3(n-1)! + 3(n-2)! - (n-3)!$$

Permutationen, in welchen keines der drei Elemente i, k, l an seinem Platze steht.

Diese Induktion zeigt, daß es

$$\varphi_n(r) = n! - \binom{r}{1}(n-1)! + \binom{r}{2}(n-2)! - \cdots + (-1)^r \binom{r}{r}(n-r)! \quad (1)$$

Permutationen gibt, in welchen keines von den r bezeichneten Elementen den ihm gebührenden Platz einnimmt.

Mithin ist die verlangte Wahrscheinlichkeit:

$$\frac{\varphi_n(r)}{m} = 1 - \binom{r}{1}\frac{1}{n} + \binom{r}{2}\frac{1}{n(n-1)} - \cdots + (-1)^r \binom{r}{r}\frac{1}{n(n-1)\cdots(n-r+1)}. \quad (2)$$

Von besonderem Interesse ist die Wahrscheinlichkeit p, daß *überhaupt keine* der Kugeln ihrem Range entsprechend erscheint; sie ergibt sich aus (2), wenn $r = n$ gesetzt wird, und ist daher

$$p = 1 - \frac{1}{1} + \frac{1}{1 \cdot 2} - \frac{1}{1 \cdot 2 \cdot 3} + \cdots + (-1)^n \frac{1}{1 \cdot 2 \cdots n}. \tag{3}$$

Daraus folgt als Wahrscheinlichkeit, daß *mindestens eine* Kugel ihrer Nummer entsprechend gezogen wird:

$$q = 1 - p = 1 - \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} - \cdots + (-1)^{n-1} \frac{1}{1 \cdot 2 \cdots n}. \tag{4}$$

Diese Aufgabe liegt einem Glücksspiel zugrunde, welches unter dem Namen „Jeu du Treize" zum erstenmal von Montmort[1]) der mathematischen Behandlung zugeführt wurde, an der sich Nicolaus Bernoulli[2]) hervorragend beteiligt hat. Das Spiel besteht in folgendem: 13 mit den Nummern 1 bis 13 beschriebene Karten werden aus einer Urne, in der sie vorher durcheinander geworfen worden sind, sukzessive herausgezogen; das Spiel gilt als gewonnen, wenn mindestens eine Karte ihrer Nummer entsprechend gezogen wurde; wie groß ist die Wahrscheinlichkeit hierfür? Dem Wesen nach gleichbedeutend hiermit ist das *Rencontrespiel*: Von zwei Spielern hat jeder ein vollständiges Spiel von 32 Blättern in der Hand; beide legen immer je ein Blatt auf: treffen gleiche Blätter zusammen, so hat sich ein Rencontre ergeben; der eine Spieler wettet auf den Eintritt, der andere auf das Ausbleiben eines Rencontres; welches sind ihre Chancen, zu gewinnen?

Dem Problem blieb fortan das mathematische Interesse zugewendet. Euler[3]) bemerkte zuerst die Beziehung, in welcher die Lösung zur Zahl e steht; die Summe (3), welche p darstellt, geht nämlich für $n = \infty$ in die Reihe über, welche e^{-1} definiert, so daß dann

$$\begin{aligned} p &= e^{-1} &&= 0{,}36787944 \cdots \\ q &= 1 - e^{-1} &&= 0{,}63212055 \cdots \end{aligned} \tag{4}$$

ist; wegen der raschen Konvergenz der betreffenden Reihe kommen die Lösungen schon für mäßig große n diesen Werten sehr nahe; so ergibt sich für $n = 13$, also für das „Jeu du Treize"[4]), in strenger Rechnung:

$$p = 0{,}367879441 \cdots$$

---

1) P. de Montmort, Essai d'Analyse sur les Jeux de Hasards (1708, 1714) p. 130 ff. 2) Ibid., p. 301—302.

3) Calcul de la probabilité dans le Jeu de Rencontre. Hist. Acad. Berlin pour 1751.

4) In der englischen Literatur „game of Treize".

auf acht Dezimalen übereinstimmend mit e^{-1}. Um so mehr können die Werte (4) für das Rencontre beibehalten werden.

Unabhängig von andern scheint Lambert[1]) auf ein ähnliches Problem gekommen zu sein, das so lautet: Jemand hat n Briefe und die zugehörigen Kuverts geschrieben; er steckt die Briefe ohne Wahl in die Kuverts; wie groß ist die Wahrscheinlichkeit, daß alle Briefe oder eine bezeichnete Anzahl derselben in die richtigen Kuverts kommen?

Lampe[2]) hat eine von Dirichlet und Kummer stammende Fassung der Aufgabe und eine Ableitung der oben mit $\varphi_n(n)$ bezeichneten Zahl mitgeteilt. Die Aufgabe ist so gestellt: Gegeben sind n Elemente auf n Plätzen; wie groß ist die Wahrscheinlichkeit, daß bei einer neuen willkürlichen Verteilung der n Elemente auf dieselben n Plätze kein Element seinen früheren Platz wieder erhalte?

Catalan[3]) hat das Problem in folgender Modifikation behandelt: Eine Urne enthält n mit den Buchstaben $a, b, c, \cdots$ markierte Kugeln; man zieht die Kugeln sukzessive einzeln heraus und legt sie dann wieder in die leere Urne zurück. Wie groß ist die Wahrscheinlichkeit, daß in zwei aufeinanderfolgenden Ziehungen m Kugeln an gleicher Stelle erscheinen?

Da die erste wie die zweite Folge der Buchstaben $n!$ verschiedene Anordnungen haben kann, so gibt es für zwei Ziehungen $n!^2$ mögliche Fälle.

Jede der $n!$ Anordnungen der ersten Ziehung kann mit der zweiten Ziehung auf $\binom{n}{m}$ verschiedene Arten Rencontres von m Buchstaben ergeben, und da jedesmal die $n-m$ übrigen Buchstaben kein Rencontre aufweisen dürfen, was wieder auf $\varphi_{n-m}(n-m)$ Arten geschehen kann, so gibt es der günstigen Fälle

$$n!\binom{n}{m}\varphi_{n-m}(n-m) = \frac{n!^2}{m!\,(n-m)!}\Big((n-m)! - \binom{n-m}{1}(n-m-1)! + \binom{n-m}{2}(n-m-2)! - \cdots + (-1)^{n-m}\Big);$$

folglich ist die verlangte Wahrscheinlichkeit gleich

$$\frac{1}{1\cdot 2\cdots m}\left(1 - \frac{1}{1} + \frac{1}{1\cdot 2} - \frac{1}{1\cdot 2\cdot 3} + \cdots + \frac{(-1)^{n-m}}{1\cdot 2\cdots(n-m)}\right).$$

---

1) Examen d'une espèce de superstition ramenée au calcul des probabilités. Nouv. Mém. Berlin 1798.

2) Über eine Aufgabe der Wahrscheinlichkeitsrechnung. Grunert, Arch. 70 (1884).

3) Solution d'un problème de probabilité, relatif au Jeu de Rencontre, Journ. Liouv, 2 (1837).

Laplace[1]) hat das dem Rencontrespiel zugrunde liegende Problem in verallgemeinerter Form behandelt.

27. Beispiel XV. *Eine Lotterie besteht aus n Nummern; in jeder Ziehung werden r Nummern gezogen. Jemand hat auf s ($\leq r$) Nummern gesetzt. Wie groß ist die Wahrscheinlichkeit, a) daß die Nummern an beliebiger Stelle und in beliebiger Ordnung; b) an beliebiger Stelle, aber in gegebener Ordnung; c) an erster Stelle in beliebiger Ordnung; d) an erster Stelle in gegebener Ordnung erscheinen?*

a) Die möglichen Fälle sind durch die Kombinationen o. W. der n Nummern zu je r dargestellt, ihre Anzahl ist also:

$$m_1 = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Scheidet man die s besetzten Nummern aus, bildet aus den $n-s$ übrigen Kombinationen zu je $r-s$ Nummern, fügt zu jeder die besetzten Nummern hinzu, so ergeben sich die günstigen Fälle, deren Anzahl somit

$$g_1 = \binom{n-s}{r-s} = \frac{(n-s)!}{(r-s)!\,(n-r)!}.$$

ist.

Die Wahrscheinlichkeit ist hiernach:

$$p_1 = \frac{(n-s)!\,r!}{n!\,(r-s)!} = \frac{r(r-1)\cdots(r-s+1)}{n(n-1)\cdots(n-s+1)}. \tag{1}$$

Bei $n = 90$ und $r = 5$ ergibt sich:

$$\begin{array}{lll} \text{für } s = 1 & p_1 = \frac{1}{18}. \\ \text{„} \quad \text{„} = 2 & \text{„} = \frac{1}{400{,}5} \\ \text{„} \quad \text{„} = 3 & \text{„} = \frac{1}{11748} \\ \text{„} \quad \text{„} = 4 & \text{„} = \frac{1}{511038} \\ \text{„} \quad \text{„} = 5 & \text{„} = \frac{1}{43949268}. \end{array}$$

b) Permutiert man in jeder der $\binom{n}{r}$ Kombinationen, welche die möglichen Fälle des vorigen Ereignisses gebildet haben, s Nummern, und zwar so, daß in jenen Kombinationen, in welchen die *s besetzten* Nummern vorkommen, gerade diese der Permutation unterworfen werden, so erhält man die möglichen Fälle für die zweite Fragestellung; ihre Anzahl ist sonach

1) Théorie anal. d. prob., Art. 9.

$$m_2 = \binom{n}{r} s!.$$

Die Anzahl der günstigen Fälle ist dieselbe wie vorhin, indem zur Gewinnung dieser Fälle jeder der $\binom{n-s}{r-s}$ Kombinationen der nicht besetzten Nummern die s besetzten Nummern in der gegebenen Ordnung anzufügen sind. Folglich ist die entsprechende Wahrscheinlichkeit:

$$p_2 = \frac{1}{1\cdot 2 \cdots s} \frac{r(r-1)\cdots(r-s+1)}{n(n-1)\cdots(n-s+1)}; \tag{2}$$

sie ist $s!$ mal kleiner als vorige.

c) Die Wahrscheinlichkeit, daß die s Nummern an erster Stelle in beliebiger Ordnung erscheinen, ist dieselbe, welche sich ergeben würde, wenn die Ziehungen nur aus s Nummern bestünden und s bezeichnete Nummern in beliebiger Ordnung erscheinen sollten; sie geht also aus p_1 durch die Substitution $r = s$ hervor und ist

$$p_3 = \frac{s(s-1)\cdots 1}{n(n-1)\cdots(n-s+1)}. \tag{3}$$

d) Ist auch noch die Ordnung bezeichnet, so stellt sich dasselbe Verhältnis wie zwischen p_2 und p_1 ein, so daß also in dem vierten Falle die Wahrscheinlichkeit

$$p_4 = \frac{1}{n(n-1)\cdots(n-s+1)} \tag{4}$$

gilt.

Für $n = 90$ und

$$s = 1 \text{ ist } p_4 = \frac{1}{90}$$

$$\text{„} = 2 \text{ „ „} = \frac{1}{8010}$$

$$\text{„} = 3 \text{ „ „} = \frac{1}{704880}$$

$$\text{„} = 4 \text{ „ „} = \frac{1}{61324560}$$

$$\text{„} = 5 \text{ „ „} = \frac{1}{5273912160}.$$

§ 3. Indirekte Methoden der Wahrscheinlichkeitsbestimmung.

28. Vorbemerkung. Wie in jedem Gebiet der Mathematik das Bestreben dahin gerichtet ist, die Lösung komplizierterer Probleme auf eine Reihe von Grundaufgaben zurückzuführen, die dann in mannigfacher Kombination zur Anwendung kommen, so sind im Laufe der

Zeit auch in der Wahrscheinlichkeitsrechnung Grundregeln und Methoden ausgebildet worden, mit deren Hilfe verwickeltere Aufgaben, bei denen das direkte Verfahren an der Schwierigkeit und Umständlichkeit der Aufstellung und Zählung der möglichen und günstigen Fälle scheitern würde, in einfacherer Weise gelöst werden können. Es sind einige wenige aus der Wahrscheinlichkeitsdefinition hervorgehende Sätze, die bei den indirekten Methoden beständig zur Anwendung kommen. So einfach diese Sätze an sich sind, so erfordert doch die Analyse des Problems, die ihrer Anwendung vorausgehen muß, scharfes Urteilen und stete Rücksichtnahme auf die Voraussetzungen, welche den Sätzen zugrunde liegen.

Der Aufstellung dieser Sätze sollen einige Betrachtungen allgemeiner Natur vorangestellt werden.

29. Absolute und relative Wahrscheinlichkeit. Abhängigkeit von Ereignissen. Es gibt vielfach Fragestellungen, bei welchen dem Ereignis, um dessen Wahrscheinlichkeit es sich handelt, gewisse Bedingungen auferlegt werden, von denen das Maß seiner Wahrscheinlichkeit abhängt, indem durch die Bedingungen eine Auslese unter den möglichen Fällen getroffen wird nach bestimmten ihnen zukommenden Qualitäten. Eine unter solchen Gesichtspunkten gerechnete Wahrscheinlichkeit bezeichnet man als eine *relative* im Gegensatze zur *absoluten*, bei der die ganze Urteilsmaterie ohne jede weitere Bedingung zur Grundlage genommen wird.

Das Verhältnis dieser beiden Wahrscheinlichkeitsbestimmungen kann an dem folgenden typischen Schema klargelegt werden.

In einer Urne befinden sich Kugeln, die durch zwei Qualitäten gekennzeichnet sind: durch eine aufgeschriebene Nummer, 0 oder 1, und durch eine Farbe, rot (r) oder schwarz (s); r_0 rote Kugeln tragen die Nummer 0, r_1 die Nummer 1; s_0 schwarze Kugeln die Nummer 0, s_1 die Nummer 1.

Wird die Frage nach der Wahrscheinlichkeit gerichtet, eine Kugel mit der Nummer 0 schlechtweg zu ziehen, so handelt es sich um eine absolute Wahrscheinlichkeit, die mit $\mathfrak{W}(0)$ bezeichnet werden möge; ihr Wert ist

$$\mathfrak{W}(0) = \frac{r_0 + s_0}{r_0 + s_0 + r_1 + s_1}.$$

Wird hingegen um die Wahrscheinlichkeit gefragt, die gezogene Kugel trage die Nummer 0 unter der Voraussetzung, daß eine rote Kugel erscheint, so hat man es mit einer relativen Wahrscheinlichkeit zu tun, zu deren Bezeichnung das Symbol $\mathfrak{W}_r(0)$ verwendet werden möge; es ist

$$\mathfrak{W}_r(0) = \frac{r_0}{r_0 + r_1}.$$

Umgekehrt kann auch gefragt werden, mit welcher Wahrscheinlichkeit eine rote Kugel zu erwarten ist unter der Voraussetzung, daß eine mit 0 bezeichnete Kugel gezogen wird; diese relative Wahrscheinlichkeit ist sinngemäß mit $\mathfrak{W}_0(r)$ zu bezeichnen und hat im vorliegenden Falle den Wert $\frac{r_0}{r_0 + s_0}$.

Um eine andere Sachlage vorzuführen, bei der zwischen absoluter und relativer Wahrscheinlichkeit unterschieden werden kann, stelle man sich vor, daß mehrere äußerlich nicht unterschiedene Urnen $U_1, U_2, \cdots U_n$ mit weißen und schwarzen Kugeln in gleicher Gesamtzahl c, aber in verschiedenem Mengenverhältnis der Farben gefüllt seien, so zwar, daß sie der Reihe nach $a_1, a_2, \cdots a_n$ weiße Kugeln enthalten.

Stellt man die offenkundig zulässige Frage nach der Wahrscheinlichkeit, daß ein Zug aus einer der Urnen eine weiße Kugel ergebe, so ist es eine absolute Wahrscheinlichkeit, $\mathfrak{W}(w)$, die man in der Weise bestimmen kann, daß man die Inhalte der Urnen in *einer* Urne vereinigt; denn nach dieser Vereinigung werden die aus verschiedenen Urnen stammenden Kugeln wieder gleichmögliche Fälle darstellen, was nicht stattfände, wenn die Urnen mit ungleichen Mengen von Kugeln gefüllt wären; es ist also

$$\mathfrak{W}(w) = \frac{a_1 + a_2 + \cdots + a_n}{nc}.$$

Daneben kann man so viele relative Wahrscheinlichkeiten eines weißen Zuges unterscheiden, als es Urnen gibt; $\mathfrak{W}_{U_i}(w)$ bedeutet die Wahrscheinlichkeit eines weißen Zuges unter der Voraussetzung, daß der Zug aus der Urne U_i erfolgt; ihr Betrag ist $\frac{a_i}{c}$.

Aber auch die Frage hat Berechtigung, mit welcher Wahrscheinlichkeit anzunehmen ist, daß der Zug aus der Urne U_i erfolgt sei, wenn vorausgesetzt wird, daß die gezogene Kugel weiß war. Da unter dieser Voraussetzung in der Urne, die die Inhalte von $U_1, U_2, \cdots U_n$ vereinigt, nur die weißen Kugeln in Betracht kommen, so ist diese Wahrscheinlichkeit, die mit $\mathfrak{W}_w(U_i)$ zu bezeichnen sein wird, gleich

$$\frac{a_i}{a_1 + a_2 + \cdots + a_n}.$$

Man wird also allgemein unter der *relativen Wahrscheinlichkeit eines Ereignisses E in bezug auf ein anderes F*, symbolisch unter $\mathfrak{W}_F(E)$, die Wahrscheinlichkeit von E verstehen, gerechnet unter der Voraussetzung, daß F verwirklicht sei. Dieser Wahrscheinlichkeit kann die inverse $\mathfrak{W}_E(F)$ gegenübergestellt werden, die sich auf das Eintreffen von F unter Voraussetzung des Bestandes von E bezieht.

Von dieser Begriffsbildung kann mit Vorteil Gebrauch gemacht werden, um dem, was man in der Wahrscheinlichkeitsrechnung als Unabhängigkeit und als Abhängigkeit zweier Ereignisse bezeichnet, einen prägnanten Ausdruck zu geben.

Angenommen, zwei Urnen U_1, U_2 enthielten c_1, c_2 Kugeln, darunter seien a_1, a_2 weiß, die andern schwarz; eine dritte Urne U enthalte unter c Kugeln a rote und $c-a$ blaue. Der Zug einer roten Kugel werde mit F, der einer blauen mit F, endlich der einer weißen mit E bezeichnet.

Man zieht aus U eine Kugel, und je nachdem sie rot oder blau ist, erfolgt der zweite Zug aus U_1 oder U_2; gefragt wird um die Wahrscheinlichkeit einer weißen Kugel.

Auf diese Frage gibt es zwei verschiedene Antworten: Liefert der erste Zug rot, so ist die gefragte Wahrscheinlichkeit $\frac{a_1}{c_1}$; liefert er blau, so ist sie $\frac{a_2}{c_2}$; nur wenn $\frac{a_1}{c_1} = \frac{a_2}{c_2}$, ist es gleichgültig, wie der erste Zug ausgefallen ist. Im ersten Falle sagt man, das Ereignis E sei von F *abhängig*, weil ihm eine verschiedene Wahrscheinlichkeit zukommt, je nachdem F eintritt oder nicht eintritt; im zweiten Falle nennt man E von F *unabhängig*, weil die Wahrscheinlichkeit von E dieselbe bleibt, ob F eingetroffen ist oder nicht.

Das Merkmal für die *Abhängigkeit* eines Ereignisses E von einem andern F ist in der Beziehung $\mathfrak{W}_F(E) \neq \mathfrak{W}_{\bar{F}}(E)$ enthalten, das der *Unabhängigkeit* in dem Ansatze $\mathfrak{W}_F(E) = \mathfrak{W}_{\bar{F}}(E)$.

Wenn auch hier die Erklärung an zeitlich aufeinanderfolgenden Ereignissen geschah, so ist doch die zeitliche Sukzession kein notwendiges Attribut der Abhängigkeit. Ferner ist die Abhängigkeit, die hier als eine einseitige dargestellt wurde, in manchen Fällen gegenseitig, indem neben $\mathfrak{W}_F(E) \neq \mathfrak{W}_{\bar{F}}(E)$ auch $\mathfrak{W}_E(F) \neq \mathfrak{W}_{\bar{E}}(F)$ besteht.

Ein besonderer Fall der Abhängigkeit zweier Ereignisse E, F ist ihre *gegenseitige Ausschließung*, darin bestehend, daß das Eintreffen des einen von beiden die Existenz des andern unmöglich macht. Dieser Sachverhalt drückt sich dadurch aus, daß $\mathfrak{W}_F(E) = \mathfrak{W}_E(F) = 0$ ist.

Hat die Voraussetzung, daß F stattfindet, die Existenz von E zur *notwendigen* Folge, so ist $\mathfrak{W}_F(E) = 1$; dies also ist der symbolische Ausdruck eines hypothetischen Urteils.[1])

---

1) Auf die wichtige Rolle, die dem Begriff der relativen Wahrscheinlichkeit bei der Grundlegung der Wahrscheinlichkeitsrechnung zugewiesen werden kann, hat F. Hausdorff, Berichte der mathem.-phys. Klasse der Sächs. Ges. d. Wissenschaften, Leipzig 1901, p. 152—164 hingewiesen. In der Tat gestattet dieser Begriff eine klare Formulierung und eine übersichtliche symbolische Darstellung vieler der nachfolgenden Sätze.

30. Vollständige Wahrscheinlichkeit. Die einem Ereignis E günstigen Fälle, g an Zahl, mögen sich von irgend einem Gesichtspunkte aus in Gruppen von $g_1, g_2, \cdots g_r$ Fällen sondern lassen derart, daß jeder der g Fälle einer und nur einer dieser Gruppen angehört; dann ist

$$g = g_1 + g_2 + \cdots + g_r.$$

Die Verwirklichung eines Falles aus der ersten Gruppe werde als Ereignis E_1, die Verwirklichung eines Falles aus der zweiten Gruppe als Ereignis E_2 bezeichnet usw. Dann sind $E_1, E_2, \cdots E_r$ Spezialfälle, Arten oder Modalitäten des Ereignisses E. Dieses Ereignis gilt als eingetroffen, entweder wenn E_1 *oder* $E_2 \cdots$ *oder* E_r eingetroffen ist, wobei wohl darauf zu achten ist, daß es jedesmal nur auf *eine* dieser Arten in Erscheinung treten kann.

Dieser Zusammenhang zwischen E und den $E_1, E_2, \cdots E_r$ soll künftighin symbolisch durch den Ansatz

$$E = E_1 + E_2 + \cdots + E_r$$

zum Ausdruck gebracht werden, der also zu lesen wäre: E ist E_1 oder E_2 oder $\cdots E_r$ und nur eines von allen.

Ist m die Anzahl der möglichen Fälle, so ist die Wahrscheinlichkeit des Ereignisses E bestimmt durch:

$$\frac{g}{m} = \frac{g_1}{m} + \frac{g_2}{m} + \cdots + \frac{g_r}{m};$$

$\frac{g_1}{m}$ ist aber die Wahrscheinlichkeit von E_1, $\frac{g_2}{m}$ die Wahrscheinlichkeit von E_2 usw.; demnach ist

$$\mathfrak{W}(E) = \mathfrak{W}(E_1 + E_2 + \cdots + E_r) = \mathfrak{W}(E_1) + \mathfrak{W}(E_2) + \cdots + \mathfrak{W}(E_r). \quad (1)$$

Dies gibt den Satz: *Wenn ein Ereignis auf mehrere einander ausschließende Arten eintreffen kann, so ist seine Wahrscheinlichkeit gleich der Summe der den einzelnen Arten des Eintreffens zukommenden Wahrscheinlichkeiten.*

Die Wahrscheinlichkeit $\mathfrak{W}(E)$ wird im Gegensatze zu den Wahrscheinlichkeiten $\mathfrak{W}(E_1), \mathfrak{W}(E_2) \cdots$, die man als *partiale* Wahrscheinlichkeiten bezeichnet, die *vollständige* oder *totale Wahrscheinlichkeit* des Ereignisses E genannt.

Man nennt die in (1) ausgesprochene Regel auch den *Additionssatz* der Wahrscheinlichkeitsrechnung.

Seine einfachste Verdeutlichung findet er in dem folgenden Beispiel. Eine Urne enthalte neben a weißen Kugeln b rote, c blaue und d gelbe Kugeln. Besteht das erwartete Ereignis in dem Erscheinen

einer *farbigen* Kugel bei Ausführung eines Zuges, so kann es auf drei verschiedene einander ausschließende Arten zustande kommen, entweder indem eine rote oder eine blaue oder eine gelbe Kugel gezogen wird; diese Arten haben der Reihe nach die Wahrscheinlichkeit $\frac{b}{m}$, $\frac{c}{m}$, $\frac{d}{m}$, wenn $m = a + b + c + d$ gesetzt wird; das Erscheinen einer farbigen Kugel überhaupt besitzt aber die Wahrscheinlichkeit $\frac{b+c+d}{m}$, und diese ist die Summe jener Einzel- oder Partialwahrscheinlichkeiten.

31. Zusammengesetzte Wahrscheinlichkeit. Ein Ereignis E gelte dann als eingetroffen, wenn sowohl das Ereignis E_1 als auch das Ereignis $E_2 \cdots$, als auch das Ereignis E_r eingetreten ist, mit andern Worten: es bestehe in dem Zusammentreffen der Ereignisse $E_1, E_2, \cdots E_r$.

Dieser Zusammenhang zwischen E und den E_1, $E_2 \cdots E_r$ wird in der Folge symbolisch durch

$$E = E_1 E_2 \cdots E_r$$

angedeutet werden, was also zu lesen ist: E ist eingetroffen, wenn E_1 und E_2 und $\cdots E_r$ existiert.

Von den $m_1, m_2, \cdots m_r$ möglichen Fällen, welche diesen Ereignissen zugrunde liegen, muß je einer eintreten, und da sich jeder Fall der ersten Gruppe mit jedem der zweiten Gruppe verbinden, jede dieser Verbindungen mit jedem Falle der dritten Gruppe sich vereinigen kann usw., so gibt es

$$m = m_1 m_2 \cdots m_r$$

Verbindungen, die in bezug auf das Ereignis E als mögliche, insbesondere auch als gleichmögliche Fälle aufzufassen sind, wenn die m_1 und die $m_2 \cdots$ und m_r Fälle untereinander gleichberechtigt waren.

Ein dem Ereignis E günstiger Fall ergibt sich nur dann, wenn je ein günstiger Fall jedes der Ereignisse $E_1, E_2, \cdots E_r$ eintrifft, und das kann auf

$$g = g_1 g_2 \cdots g_r$$

Arten gechehen, wenn dem E_1 g_1 günstige Fälle, dem E_2 g_2 günstige Fälle usw. zukommen. Hiernach ist die Wahrscheinlichkeit des Ereignisses E bestimmt durch:

$$\frac{g}{m} = \frac{g_1}{m_1} \frac{g_2}{m_2} \cdots \frac{g_r}{m_r};$$

nun ist aber $\frac{g_1}{m_1}$ die Wahrscheinlichkeit des Ereignisses E_1 an sich, $\frac{g_2}{m_2}$ die Wahrscheinlichkeit von E_2 usw., demnach

$$\mathfrak{W}(E) = \mathfrak{W}(E_1 E_2 \cdots E_r) = \mathfrak{W}(E_1)\mathfrak{W}(E_2) \cdots \mathfrak{W}(E_r). \tag{2}$$

Bei der vorstehenden Betrachtung ist *Unabhängigkeit* der Ereignisse $E_1, E_2, \cdots$ vorausgesetzt, so daß $\mathfrak{W}_{E_i}(E_k) = \mathfrak{W}_{\bar{E}_i}(E_k)$ für jede Ambe aus den Zeigern $1, 2 \cdots r$ stattfindet. Die Gleichung (2) gibt hiernach den Satz: *Die Wahrscheinlichkeit für das Zusammentreffen mehrerer voneinander unabhängiger Ereignisse ist das Produkt der Wahrscheinlichkeiten dieser Ereignisse.*

Die Wahrscheinlichkeit $\mathfrak{W}(E)$ wird im Gegensatze zu den Wahrscheinlichkeiten $\mathfrak{W}(E_1), \mathfrak{W}(E_2), \cdots \mathfrak{W}(E_r)$, die man *einfache* Wahrscheinlichkeiten heißt, eine *zusammengesetzte Wahrscheinlichkeit* genannt. Auch das Ereignis E wird als ein aus den *einfachen* Ereignissen $E_1, E_2 \cdots E_r$ zusammengesetztes Ereignis bezeichnet.

Die durch (2) ausgedrückte Regel wird der *Multiplikationssatz* der Wahrscheinlichkeitsrechnung genannt.

Ob das Zusammentreffen ein gleichzeitiges oder ein sukzessives ist, bleibt für die Erwartungsbildung und daher auch für die Wahrscheinlichkeitsbestimmung gleichgültig; im Falle des folgeweisen Eintreffens der einfachen Ereignisse ist auch ihre Ordnung ohne Belang; hierin liegt das wahrscheinlichkeitstheoretische Äquivalent für das kommutative Multiplikationsgesetz.

Wenn die Ereignisse $E_2, E_3, \cdots E_r$ mit dem ersten, E_1, übereinstimmen, so besteht das Ereignis E, das nun mit E_1^r zu bezeichnen sein wird, entweder in dem Zusammentreffen von r gleichwahrscheinlichen oder in der r-maligen Wiederholung eines und desselben Ereignisses, und seine Wahrscheinlichkeit ist zufolge (2):

$$\mathfrak{W}(E) = \mathfrak{W}(E_1^r) = \mathfrak{W}(E_1)^r. \tag{3}$$

Es ist also die Wahrscheinlichkeit der r-maligen Wiederholung eines und desselben Ereignisses gleich der r-ten Potenz seiner Wahrscheinlichkeit.

Die Multiplikationsregel erfährt eine das Wesen der Sache berührende Änderung, wenn die einfachen Ereignisse $E_1, E_2, \cdots$, aus denen sich E zusammensetzt, *voneinander abhängig* sind in der Weise, daß die Wahrscheinlichkeit eines Ereignisses der Reihe abhängt von dem Eintreffen oder Nichteintreffen der vorausgehenden. In diesem Falle kommen vom zweiten Ereignis angefangen nicht mehr deren absolute, sondern die *relativen* Wahrscheinlichkeiten in Rechnung, gebildet unter der Voraussetzung, daß alle vorangehenden Ereignisse eingetroffen sind, wie es die Existenz von E erfordert; es tritt also an die Stelle von $\mathfrak{W}(E_2)$ jetzt $\mathfrak{W}_{E_1}(E_2)$, an die Stelle von $\mathfrak{W}(E_3)$ nunmehr $\mathfrak{W}_{E_1 E_2}(E_3)$, endlich kommt statt $\mathfrak{W}(E_r)$ zu setzen $\mathfrak{W}_{E_1 E_2 \cdots E_{r-1}}(E_r)$, so daß bei der gegenwärtigen Sachlage

$$\mathfrak{W}(E) = \mathfrak{W}(E_1)\mathfrak{W}_{E_1}(E_2)\mathfrak{W}_{E_1 E_2}(E_3) \cdots \mathfrak{W}_{E_1 E_2 \cdots E_{r-1}}(E_r). \tag{4}$$

In Worten: *Die Wahrscheinlichkeit für das Zusammentreffen mehrerer voneinander abhängiger Ereignisse ist gleich dem Produkt ihrer Wahrscheinlichkeiten, die Wahrscheinlichkeit jedes einzelnen Ereignisses unter der Voraussetzung gerechnet, daß die ihm in der Sukzession vorangehenden eingetroffen seien.*

Nach dem Gange der Herleitung möchte es scheinen, als ob im Falle der Abhängigkeit eine bestimmte Ordnung der einfachen Ereignisse vorausgesetzt werden müßte. Bei Anwendungen auf die Wirklichkeit ist allerdings häufig eine zeitliche Sukzession durch die Natur der Sache bestimmt; es gibt aber auch Materien, bei denen das $\mathfrak{W}(E)$ der Formel (4) ebenso unabhängig ist von der Ordnung der $E_1, E_2, \cdots$ wie dies bei dem $\mathfrak{W}(E)$ der Formel (2) immer stattfindet.

Zur Erläuterung der beiden Modalitäten zusammengesetzter Ereignisse mögen die folgenden einfachen Beispiele dienen.

Aus jedem von zwei Kartenspielen zu 32 Blättern wird eine Karte gezogen; wie groß ist die Wahrscheinlichkeit, daß man zwei Könige zieht? Die Unabhängigkeit der beiden Ereignisse, die zusammentreffen sollen, ist hier offenkundig: was der Zug aus dem einen Spiele bringen möge, die Wahrscheinlichkeit, aus dem andern einen König zu ziehen, bleibt $\frac{1}{8}\cdot$ Die Wahrscheinlichkeit des zusammengesetzten Ereignisses ist $\frac{1}{8}\cdot\frac{1}{8}=\frac{1}{64}\cdot$

Nun sollen beide Karten aus demselben Spiele nacheinander gezogen werden; welches ist die Wahrscheinlichkeit, daß es zwei Könige seien? Ist das erstgezogene Blatt ein König — und nur unter dieser Voraussetzung kann das bezeichnete Ereignis eintreten —, so gibt es für das zweite Ereignis nur mehr 31 mögliche und darunter nur noch 3 günstige Fälle. Demnach ist jetzt die Wahrscheinlichkeit des zusammengesetzten Ereignisses $\frac{1}{8}\cdot\frac{3}{31}=\frac{3}{248}\cdot$

Dieser Wert gilt indessen auch dann, wenn die Karten statt nacheinander auf einmal gezogen werden. Denn, faßt man das Ereignis dann als ein einfaches auf, so hat man, entsprechend den zweiblättrigen Kombinationen, $\frac{32\cdot 31}{1\cdot 2}$ mögliche und, entsprechend den zweiblättrigen Kombinationen der Könige, $\frac{4\cdot 3}{1\cdot 2}$ günstige Fälle; folglich ist die Wahrscheinlichkeit $\frac{4\cdot 3}{32\cdot 31}=\frac{3}{248}$ wie oben.

32. Beispiel XVI. *Es ist die Wahrscheinlichkeit zu bestimmen, mit zwei Würfeln eine ungerade Summe zu werfen.*

Nach Nr. 14 haben die ungeraden Summen 3, 5, 7, 9, 11, welche als Arten des Eintreffens des bezeichneten Ereignisses aufzufassen sind, die Wahrscheinlichkeiten

$$p_3 = p_{11} = \frac{1}{18}, \quad p_5 = p_9 = \frac{1}{9}, \quad p_7 = \frac{1}{6};$$

daraus ergibt sich die vollständige Wahrscheinlichkeit einer ungeraden Summe

$$p = \frac{1}{9} + \frac{2}{9} + \frac{1}{6} = \frac{1}{2}.$$

33. Beispiel XVII. *Eine Urne U enthält a weiße und b schwarze, eine zweite ihr äußerlich gleiche Urne, U', a' weiße und b' schwarze Kugeln; man zieht aus einer von beiden eine Kugel heraus; wie groß ist die Wahrscheinlichkeit, daß sie weiß (w) sei?*

Angenommen, es sei $a = 3$, $b = 4$; $a' = 4$, $b' = 3$. Es könnte hier die Meinung entstehen, daß es sich um zwei einander ausschließende Arten eines Ereignisses und daher um die Berechnung einer vollständigen Wahrscheinlichkeit handle; der Ansatz

$$\frac{3}{7} + \frac{4}{7} = 1$$

würde aber diese Meinung sofort als falsch erkennen lassen, da es gewiß nicht notwendig ist, eine weiße Kugel zu ziehen. Es muß vielmehr beachtet werden, daß jede Entstehungsart selbst als zufällig ihre Wahrscheinlichkeit, und zwar $\frac{1}{2}$, hat, weil es ebenso leicht möglich ist, in die erste wie in die zweite Urne zu greifen; richtig ist also die Bestimmung

$$\frac{1}{2} \cdot \frac{3}{7} + \frac{1}{2} \cdot \frac{4}{7} = \frac{1}{2}.$$

Allgemein ist $\frac{1}{2}$ die Wahrscheinlichkeit, in die erste Urne zu greifen, und $\frac{a}{a+b}$ die Wahrscheinlichkeit, aus ihr eine weiße Kugel zu ziehen, $\frac{1}{2} \cdot \frac{a}{a+b}$ die Wahrscheinlichkeit für beides zusammen; ebenso ist $\frac{1}{2} \cdot \frac{a'}{a'+b'}$ die Wahrscheinlichkeit, in die zweite Urne zu greifen und eine weiße Kugel zu ziehen; die vollständige Wahrscheinlichkeit des erwarteten Ereignisses ist also

$$p = \frac{1}{2}\left(\frac{a}{a+b} + \frac{a'}{a'+b'}\right). \tag{1}$$

Man könnte auch auf den Gedanken kommen, den Inhalt beider Urnen in einer zu vereinigen und aus dieser zu ziehen. Dieser Vorgang führt aber nur bedingt zu dem richtigen Resultate; während nämlich die Kugeln, solange sie in zwei Urnen getrennt waren, gleichmögliche Fälle vorstellten, braucht dies nach der Vereinigung

nicht mehr zuzutreffen; sind nämlich die Gesamtzahlen $a+b$ und $a'+b'$ ungleich, so ist es nicht mehr gleich leicht, eine aus der ersten Urne stammende oder eine solche aus der zweiten Urne zu ergreifen. Ist aber $a+b=a'+b'$, dann hat der Ausdruck (1) gleichen Wert mit dem Ausdruck

$$\frac{a+a'}{a+b+a'+b'}, \tag{2}$$

welcher die Wahrscheinlichkeit für das Ziehen einer weißen Kugel aus der dritten Urne angibt; beide gehen in

$$\frac{a+a'}{2(a+b)}$$

über.

Aber auch dann werden (1) und (2) einander gleich und der Vorgang zulässig, wenn $\frac{a}{b}=\frac{a'}{b'}$ ist, wenn also das Mischungs*verhältnis* in beiden Urnen dasselbe ist; der gemeinsame Wert von (1) und (2) ist dann

$$\frac{a}{a+b}.$$

Den ersten dieser beiden Fälle kann man, wenn er nicht vorhanden ist, immer herbeiführen. Angenommen, es sei $a+b \neq a'+b'$ und N ein gemeinsames Vielfaches beider Zahlen, so daß

$$N=\lambda(a+b)=\lambda'(a'+b');$$

man substituiere den Urnen zwei andere, die eine mit λa weißen und λb schwarzen, die andere mit $\lambda' a'$ weißen und $\lambda' b'$ schwarzen Kugeln; da diese Urnen gleiche Kugelmengen enthalten, so kann man sie zu einer vereinigen, welche somit $\lambda a + \lambda' a'$ weiße und $2N$ Kugeln überhaupt enthält; die Wahrscheinlichkeit, eine weiße Kugel zu ziehen, ist

$$\frac{\lambda a+\lambda' a'}{2N}=\frac{1}{2}\left(\frac{a}{a+b}+\frac{a'}{a'+b'}\right),$$

in Übereinstimmung mit (1).

Am durchsichtigsten gestaltet sich die Lösung in symbolischer Darstellung:

$$\begin{aligned}\mathfrak{W}(w) &= \mathfrak{W}(Uw)+\mathfrak{W}(U'w)\\ &= \mathfrak{W}(U)\mathfrak{W}_U(w)+\mathfrak{W}(U')\mathfrak{W}_{U'}(w).\end{aligned}$$

34. Beispiel XVIII. *In einer Urne liegen drei mit den Nummern 1, 2, 3 markierte Kugeln; man zieht zweimal nacheinander eine Kugel*

heraus, legt jedoch die erstgezogene vor der zweiten Ziehung zurück. Wie groß ist die Wahrscheinlichkeit, daß die höchste erschienene Nummer 2 sei?

Zur Lösung dieser Aufgabe bieten sich verschiedene Auffassungen dar.

Man kann sagen: Es darf 3 nicht erscheinen und 1 nicht zweimal. Daß 3 in keinem Zuge erscheint, hat die Wahrscheinlichkeit $\frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$; daß 1 nicht zweimal erscheint, hat als Gegensatz dazu, daß es zweimal erscheint, die Wahrscheinlichkeit $1 - \left(\frac{1}{3}\right)^2 = \frac{8}{9}$. Es wäre aber unrichtig, zu schließen, das erwartete Ereignis habe die Wahrscheinlichkeit $\frac{4}{9} \cdot \frac{8}{9} = \frac{32}{81}$, weil die Ereignisse nicht unabhängig voneinander sind; denn unter der Voraussetzung, daß 3 nicht erscheint, ist die Wahrscheinlichkeit, daß 1 nicht zweimal erscheint, $1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}$, weil es sich nur mehr um die Nummern 1 und 2 handeln kann; folglich ist die richtig bestimmte Wahrscheinlichkeit $\frac{4}{9} \cdot \frac{3}{4} = \frac{1}{3}$.

Auch die folgende Analyse könnte vorgenommen werden: Daß 2 im ersten Zuge erscheint und 3 im zweiten Zuge nicht erscheint, hat die Wahrscheinlichkeit $\frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$; daß 3 im ersten Zuge nicht erscheint und 2 im zweiten Zuge herauskommt, hat die Wahrscheinlichkeit $\frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$. Es wäre aber wieder falsch, die Wahrscheinlichkeit des erwarteten Ereignisses durch die Summe $\frac{2}{9} + \frac{2}{9}$ darzustellen; denn das zweimalige Eintreffen von 2, für welches die Wahrscheinlichkeit $\frac{1}{9}$ besteht, ist bei dieser Schlußweise zweimal in Rechnung gebracht, muß also einmal subtrahiert werden, wodurch sich $\frac{2}{9} + \frac{2}{9} - \frac{1}{9} = \frac{1}{3}$ wie oben ergibt.

Am durchsichtigsten ist die folgende vollständige Analyse des Ereignisses; dasselbe kann nämlich auf folgende drei einander ausschließende Arten zustande kommen:

Im ersten Wurfe 2,	im zweiten 1;	Wahrsch.	$= \frac{1}{9}$
„ „ „ 2,	„ „ 2;	„	$= \frac{1}{9}$
„ „ „ 1,	„ „ 2;	„	$= \frac{1}{9}$

daher ist die vollständige Wahrscheinlichkeit $\frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$.

35. Beispiel XIX. *Wie groß ist die Wahrscheinlichkeit, eine bezeichnete Nummer in einer Ziehung der gewöhnlichen Lotterie zu treffen?*

Die Wahrscheinlichkeit, daß die Nummer im ersten Zuge erscheint, ist $\frac{1}{90}$.

Die Wahrscheinlichkeit, daß sie, wenn sie im ersten Zuge nicht erschien, im zweiten kommt, ist $\frac{1}{89}$.

Die Wahrscheinlichkeit, daß sie im dritten Zuge erscheint, wenn sie in den zwei ersten nicht kam, ist $\frac{1}{88}$.

Für den vierten und fünften Zug ergeben sich unter den analogen Voraussetzungen die Wahrscheinlichkeiten $\frac{1}{87}$ und $\frac{1}{86}$.

Dies sind die fünf verschiedenen und einander ausschließenden Arten des Erscheinens der bezeichneten Nummer; es wäre aber falsch, zu schließen, die Wahrscheinlichkeit für ihr Erscheinen überhaupt sei

$$\frac{1}{90}+\frac{1}{89}+\frac{1}{88}+\frac{1}{87}+\frac{1}{86}=0{,}05683\cdots.$$

Denn im zweiten Zuge hat die Nummer die Wahrscheinlichkeit $\frac{1}{89}$ nur dann, *wenn man weiß*, daß sie im ersten Zuge nicht erschienen ist; *vor* der Ziehung ist dies aber bloß eine Voraussetzung, deren Wahrscheinlichkeit $\frac{89}{90}$ beträgt. Ebenso hat die Nummer die Wahrscheinlichkeit $\frac{1}{88}$, im dritten Zuge zu erscheinen, nur dann, wenn man weiß, daß sie in den zwei ersten Zügen nicht erschienen ist; vor Beginn der Ziehung ist dies aber nur eine Voraussetzung, der die Wahrscheinlichkeit $\frac{89}{90}\cdot\frac{88}{89}$ zukommt usw. Der richtige Ansatz für die gesuchte Wahrscheinlichkeit ist daher

$$p=\frac{1}{90}+\frac{89}{90}\cdot\frac{1}{89}+\frac{89}{90}\cdot\frac{88}{89}\cdot\frac{1}{88}+\frac{89}{90}\cdot\frac{88}{89}\cdot\frac{87}{88}\cdot\frac{1}{87}+\frac{89}{90}\cdot\frac{88}{89}\cdot\frac{87}{88}\cdot\frac{86}{87}\cdot\frac{1}{86}$$
$$=\frac{1}{90}+\frac{1}{90}+\frac{1}{90}+\frac{1}{90}+\frac{1}{90}=\frac{1}{18}.$$

Man vergl. hierzu Nr. 27.

36. Beispiel XX. *Die Wahrscheinlichkeit zu bestimmen, daß ein Ereignis von gegebener Wahrscheinlichkeit p sich in n Realisierungen der allgemeinen Bedingungen wenigstens einmal, also überhaupt zutrage.*

Das erwartete Ereignis ist der Gegensatz dessen, daß das betreffende Ereignis sich in den n Versuchs- oder Beobachtungsfällen niemals einstellt, wofür $(1-p)^n$ die Wahrscheinlichkeit ist; demnach ist seine Wahrscheinlichkeit:

$$P = 1 - (1 - p)^n. \tag{1}$$

Sie wird mit wachsendem n immer größer und nähert sich der Einheit, wie klein auch p, d. h. wie unwahrscheinlich das betreffende Ereignis im einzelnen Falle auch sein möge.

Zu gegebenem P läßt sich die Zahl n der Realisierungen bestimmen, welche nötig sind, damit das Eintreffen des Ereignisses mit der Wahrscheinlichkeit P erwartet werden dürfe; es ist

$$n \doteq \frac{\log(1 - P)}{\log(1 - p)}. \tag{2}$$

Es liegt in der Natur der Sache, daß $P > p$ sein muß. (Man vergl. hierzu Nr. 23 und 24.)

37. Beispiel XXI. *Eine Lotterie besteht aus n Nummern; r davon werden in jeder Ziehung gezogen. Es ist die Wahrscheinlichkeit zu bestimmen, daß in i Ziehungen alle Nummern erscheinen.*

Die Zahl i ist von vornherein an die Bedingung $ir \gtreqless n$ gebunden; je höher sie über ihrer untern Grenze $\frac{n}{r}$ angenommen wird, um so wahrscheinlicher ist das Ereignis, von dem oben die Rede ist. So können in der gewöhnlichen Lotterie die Nummern nicht früher als nach 18 Ziehungen erschöpft sein.[1])

1) Die Wahrscheinlichkeit, daß in 18 Ziehungen alle Nummern daran kommen, oder, was dasselbe ist, daß keine Nummer sich wiederholt, ergibt sich wie folgt: Die Zahl der möglichen Fälle in *einer* Ziehung ist $\binom{90}{5}$, die Zahl der möglichen Fälle in 18 Ziehungen ist daher $\binom{90}{5}^{18}$. Welche von den $\binom{90}{5}$ Kombinationen in der ersten Ziehung auch erscheint, für die zweite Ziehung sind nur mehr 85 Nummern zulässig, aus welchen sich $\binom{85}{5}$ Kombinationen bilden lassen; für die dritte Ziehung bleiben noch 80 Nummern, und es kann eine beliebige der $\binom{80}{5}$ aus ihnen zu bildenden Kombinationen erscheinen. Setzt man diese Betrachtung fort, so kommt man zu der Zahl günstiger Fälle:

$$\binom{90}{5}\binom{85}{5}\binom{80}{5}\cdots\binom{5}{5} = \frac{90!}{5!\,85!}\cdot\frac{85!}{5!\,80!}\cdot\frac{80!}{5!\,75!}\cdot\;\cdot\frac{5!}{5!\,0!} = \frac{90!}{5!^{18}}.$$

Die gesuchte Wahrscheinlichkeit ist daher

$$\frac{90!}{\left[5!\binom{90}{5}\right]^{18}} = \frac{85!^{18}}{90!^{17}};$$

mit Benützung der Stirlingschen Formel findet man als Wert dieses Ausdrucks einen Dezimalbruch, der an der 37. Stelle die erste bedeutsame Ziffer hat, nämlich $0,\overset{36}{0}14897\cdots$

Die Wahrscheinlichkeit, daß eine bestimmte Nummer in einer Ziehung erscheint, ist $\frac{r}{n}$; die Wahrscheinlichkeit, daß sie in i Ziehungen mindestens einmal erscheint, beträgt also nach Nr. 36

$$1-\left(1-\frac{r}{n}\right)^i.$$

Dieser Ausdruck gilt für jede Nummer, die man herausgreift. Die Wahrscheinlichkeit, daß jede der n Nummern in den i Ziehungen mindestens einmal erschienen sein werde, beträgt sonach

$$P=\left[1-\left(\frac{n-r}{n}\right)^i\right]^n. \tag{1}$$

Diese Wahrscheinlichkeit deckt sich aber nicht mit der Wahrscheinlichkeit dafür, daß mit der i-ten Ziehung gerade die letzte der noch fehlenden Nummern erschienen und dadurch das Ereignis, das in Rede steht, vollendet worden ist. Für diese hat Laplace[1]) eine strenge Analyse entwickelt, welche jedoch zeigt, daß zwischen den beiden Wahrscheinlichkeiten, sofern nur $\frac{r}{n}$ ein kleiner Bruch, in jeder Ziehung also nur ein kleiner Teil der Nummern gezogen wird, ein sehr geringer Unterschied besteht.

Rechnet man für $n=90$, $r=5$ und $P=\frac{1}{2}$ die Zahl i der Ziehungen aus der Gleichung (1), wofür sich

$$\frac{\log\dfrac{\sqrt[90]{2}}{\sqrt[90]{2}-1}}{\log 90-\log 85}$$

ergibt, so erhält man $i=85{,}20\cdots$; die strenge Rechnung im Sinne der Laplaceschen Auffassung liefert $i=85{,}53\cdots$. Die numerisch sehr wenig voneinander abweichenden Resultate sind praktisch gleichbedeutend; beide besagen, daß es vorteilhaft ist, Eins gegen Eins zu wetten, daß in 86 aufeinander folgenden Ziehungen alle 90 Nummern erschienen sein werden, während die gleiche Wette auf 85 Ziehungen nicht vorteilhaft wäre.

38. Beispiel XXII. *Die Wahrscheinlichkeit zu bestimmen, daß eine Lotterieziehung aus lauter einziffrigen, aus vier einziffrigen und einer zweiziffrigen, ··· schließlich aus lauter zweiziffrigen Nummern bestehe.*

Daß im ersten Zuge eine einziffrige Nummer erscheint, hat die Wahrscheinlichkeit $\frac{9}{90}$; ist dies eingetreteten, so hat eine einziffrige

1) Théorie anal. des prob., Art. 4.

Nummer im zweiten Zuge die Wahrscheinlichkeit $\frac{8}{89}$ usw.; die Wahrscheinlichkeit einer aus lauter einziffrigen Nummern zusammengesetzten Ziehung ist daher

$$\frac{9}{90}\cdot\frac{8}{89}\cdot\frac{7}{88}\cdot\frac{6}{87}\cdot\frac{5}{86}, \text{ symbolisch } \frac{9^{5/-1}}{90^{5/-1}}.$$

Bei einer Ziehung der zweiten Gattung nehmen wir zuerst an, daß die einziffrigen Nummern vorangehen; ihre Wahrscheinlichkeit hat dann den Ausdruck:

$$\frac{9}{90}\cdot\frac{8}{89}\cdot\frac{7}{88}\cdot\frac{6}{87}\cdot\frac{81}{86};$$

für jede andere Anordnung der ein- und zweiziffrigen Nummern erhält man denselben Wert, so z. B., wenn die zweiziffrige Nummer an erster Stelle erscheinen soll:

$$\frac{81}{90}\cdot\frac{9}{89}\cdot\frac{8}{88}\cdot\frac{7}{87}\cdot\frac{6}{86},$$

was nur der Form nach von dem früheren Ausdruck verschieden ist; da es nun $\frac{5!}{4!} = 5$ verschiedene Anordnungen gibt, so ist die vollständige Wahrscheinlichkeit einer Ziehung mit bloß einer zweiziffrigen Nummer:

$$5\,\frac{9^{4/-1}\,81}{90^{5/-1}}.$$

Durch analoge Schlüsse ergeben sich für die vier noch übrigen Gattungen von Ziehungen die Wahrscheinlichkeiten:

$$10\,\frac{9^{3/-1}\,81^{2/-1}}{90^{5/-1}},\quad 10\,\frac{9^{2/-1}\,81^{3/-1}}{90^{5/-1}},\quad 5\,\frac{9\cdot 81^{4/-1}}{90^{5/-1}},\quad \frac{81^{5/-1}}{90^{5/-1}}.$$

Die Ausrechnung liefert für die sechs bezeichneten Gattungen die folgenden Wahrscheinlichkeitswerte:

0,000003
0,000232
0,006193
0,069888
0,340703
0,582981

Summe 1.

Bezüglich der ersten und letzten Gattung vgl. Nr. 19.

39. Beispiel XXIII. *Über zwei entgegengesetzte Ereignisse E, $\bar{E}$ mit den Wahrscheinlichkeiten p, q $(= 1 - p)$ werden s Versuche oder*

Beobachtungen angestellt. Wie groß ist die Wahrscheinlichkeit, daß sich das Ereignis E dabei $m(\leq s)$-mal und das Ereignis $\bar{E}$ $n\,(=s-m)$-mal in was immer für einer Reihenfolge zutragen werde?

Irgend eine Kombination der Ereignisse E, $\bar{E}$, welche s Elemente umfaßt und in der E m-mal, $\bar{E}$ n-mal vorkommt, hat zur Wahrscheinlichkeit ein Produkt von s Faktoren, wovon m gleich p und n gleich q sind, so daß $p^m q^n$ der einfachste Ausdruck dieses Produktes ist.

Kombinationen von dieser Art gibt es aber so viele, als s Elemente, worunter m gleiche einer Art und n gleiche einer andern Art vorkommen, Permutationen ergeben, d. i. $\frac{s!}{m!\,n!}$.

Da jeder solchen Anordnung die Wahrscheinlichkeit $p^m q^n$ zukommt, so ist die totale Wahrscheinlichkeit, daß die Ereignisse E, $\bar{E}$ in den Anzahlen m, n in was immer für einer Ordnung sich wiederholen werden, dargestellt durch

$$\mathfrak{W}(E^m \bar{E}^n) = \frac{s!}{m!\,n!} p^m q^n. \tag{1}$$

Dieser Ausdruck ist das allgemeine Glied der Entwicklung von $(p+q)^s$, welche lautet:

$$(p+q)^s = p^s + \binom{s}{1} p^{s-1} q + \binom{s}{2} p^{s-2} q^2 + \cdots + \binom{s}{n} p^m q^n + \cdots + q^s$$

oder mit anderer Schreibung der Koeffizienten:

$$\begin{aligned}(p+q)^s = p^s + \frac{s!}{(s-1)!\,1!} p^{s-1} q + \frac{s!}{(s-2)!\,2!} p^{s-2} q^2 + \cdots \\ + \frac{s!}{m!\,n!} p^m q^n + \cdots + q^s.\end{aligned} \tag{2}$$

Jedes Glied dieser Entwicklung hat sonach eine bestimmte wahrscheinlichkeitstheoretische Bedeutung. So stellt das erste Glied die Wahrscheinlichkeit dar, daß in allen s Versuchen das Ereignis E auftreten werde; das zweite Glied die Wahrscheinlichkeit, daß E sich $s-1$-mal wiederholen und $\bar{E}$ einmal zutragen werde, gleichgültig ob $\bar{E}$ an erster, zweiter $\cdots$ oder letzter Stelle erscheint usw.; das letzte Glied gibt die Wahrscheinlichkeit der beständigen Wiederholung von $\bar{E}$.

Weil $p + q = 1$, so ist die Summe sämtlicher Glieder der Entwicklung gleich 1, was mit der Tatsache übereinstimmt, daß von den den Gliedern entsprechenden Ereignisverbindungen nur eine eintreten kann und eine notwendig eintreten muß.

40. Beispiel XXIV. *Zwei Spieler, A und B, brechen ein Spiel in dem Augenblicke ab, wo dem ersten noch m, dem zweiten noch n Partien*

zum Gewinnen des Spieles fehlen. Wie haben sie sich in den Einsatz zu teilen, wenn p, q die bezüglichen Wahrscheinlichkeiten für das Gewinnen einer Partie sind?

Die Teilung hat billigerweise im Verhältnis der Wahrscheinlichkeiten zu erfolgen, welche die Spieler besitzen, das Spiel zu gewinnen; es kommt daher auf die Bestimmung dieser Wahrscheinlichkeiten an, die wir mit P, Q bezeichnen wollen.

Die Lösung kann nach verschiedenen Methoden erfolgen.

Man kann von dem Gedanken ausgehen, daß das Spiel sicher zur Entscheidung kommen müßte, wenn es um $m + n - 1$ Partien fortgesetzt würde; denn einer der Spieler gewänne dann sicher die ihm noch fehlenden Partien. Die Entscheidung fiele zugunsten des A aus, wenn er alle diese Partien, oder eine weniger oder zwei weniger usw., mindestens aber die ihm fehlenden m Partien gewänne. Setzt man zur Abkürzung $m + n - 1 = r$ und entwickelt $(p + q)^r$, so geben jene Glieder, in welchen der Exponent von p gleich $r, r - 1, r - 2, \cdots m$ ist, die den angeführten Modalitäten entsprechenden Wahrscheinlichkeiten, und die Summe dieser Glieder ist die vollständige Wahrscheinlichkeit P, daß A gewinnen würde; also ist

$$P = p^r + \binom{r}{1} p^{r-1} q + \binom{r}{2} p^{r-2} q^2 + \cdots + \binom{r}{n-1} p^m q^{n-1}. \quad (1)$$

Die Summe der übrigen Glieder gibt Q; da jedoch, wie wir voraussetzen, $p + q = 1$ ist, so ist auch $P + Q = 1$, daher $Q = 1 - P$.

Gegen diese Lösung könnte die Bemerkung gemacht werden, daß die Entscheidung des Spieles unter Umständen schon nach einer kleineren Anzahl von Partien fallen könnte als $m + n - 1$.

Die geringste Zahl von Partien, um welche das Spiel fortgesetzt werden müßte, damit A gewinnen könne, ist m, und zwar müßte A alle diese Partien gewinnen, wofür die Wahrscheinlichkeit

$$p^m$$

ist.

Tritt dies nicht ein, so kann A in $m + 1$ Partien gewinnen, wenn nur eine davon, jedoch nicht die letzte, an B fällt; denn sonst hätte A schon mit der m-ten Partie gewonnen gehabt; da somit für B noch m Plätze frei bleiben, so ist die Wahrscheinlichkeit für diesen Modus des Gewinnens

$$m p^m q.$$

Gewinnt A auch in $m + 1$ Partien nicht, so kann er in $m + 2$ Partien gewinnen, wenn dem B nur zwei Partien zufallen, jedoch so, daß er nicht als letzter gewinnt; denn sonst hätte A schon mit $m + 1$ oder gar mit m Partien gewonnen gehabt; da sich nun $m - 1$ Ge-

winne des A und 2 Gewinne des B auf die $m+1$ ersten Plätze auf $\frac{(m+1)!}{(m-1)!\,2!} = \frac{m(m+1)}{1\cdot 2}$ Arten verteilen können, so gewinnt A unter den gegenwärtigen Voraussetzungen mit der Wahrscheinlichkeit

$$\frac{m(m+1)}{1\cdot 2}p^m q^2.$$

Durch ähnliche Schlüsse findet man die Wahrscheinlichkeit, daß A in $m+3$ Partien gewinnt, wenn er nicht schon früher gewonnen hat, gleich

$$\frac{m(m+1)(m+2)}{1\cdot 2\cdot 3}p^m q^3.$$

So fortfahrend, kommt man schließlich zu der Anzahl $m+n-1=r$ von Partien; mit diesen *muß* aber A das Spiel gewinnen, wenn es nicht schon früher zu seinen Gunsten beendigt worden ist; da der Fall, daß B die letzte Partie gewinnt, wieder ausgeschlossen ist, und da sich die $m-1$ Partien des A und die $n-1$ Partien des B auf die $r-1$ ersten Plätze in $\frac{(r-1)!}{(m-1)!\,(n-1)!} = \frac{m(m+1)\cdots(r-1)}{1\cdot 2\cdots(n-1)}$ verschiedenen Arten verteilen können, so entspricht diesem Modus des Gewinnens die Wahrscheinlichkeit

$$\frac{m(m+1)\cdots(r-1)}{1\cdot 2\cdots(n-1)}p^m q^{n-1}.$$

Die vollständige Wahrscheinlichkeit ist sonach

$$P=p^m\left\{1+mq+\frac{m(m+1)}{1\cdot 2}q^2+\frac{m(m+1)(m+2)}{1\cdot 2\cdot 3}q^3+\cdots+\frac{m(m+1)\cdots(r-1)}{1\cdot 2\cdots(n-1)}q^{n-1}\right\}\cdot \quad (2)$$

Dieser Ausdruck ist aber wegen $p+q=1$ gleichwertig mit dem Ausdruck (1), wie man sich überzeugt, wenn man aus dem letzteren p^m heraushebt, innerhalb der Klammer p durch q ausdrückt und nach Potenzen von q ordnet.

Das vorliegende Problem nimmt in der Literatur unter den Namen „problème des partis“, „problem of points“ und „Teilungsproblem“ einen hervorragenden Platz ein. Es war eines der ersten, die gestellt und gelöst worden sind. Derselbe Chevalier de Meré, von dem in Nr. 24 die Rede war, hatte es Pascal[1]) vorgelegt, und dieser teilte es Fermat[2]) mit. Beide gaben von verschiedenen Gesichtspunkten ausgehende Lösungen, jedoch der Stellung der Aufgabe entsprechend nur für besondere Werte von m und n und unter der

---

1) Œuvres de B. Pascal (Paris, Hachette) 1858, t. II, p. 393. Der Briefwechsel zwischen Pascal und Fermat über das Problem fällt in das Jahr 1654 und findet sich in demselben Werke t. II, p. 392—407 wiedergegeben.

2) P. Fermat, Varia opera mathematica (1679).

Voraussetzung, daß für beide Spieler das Gewinnen der Partie gleich wahrscheinlich sei. An dieser letzteren Voraussetzung hielten auch Huygens[1]) und J. Bernoulli fest. Letzterer gab auf dieser Grundlage eine Tafel[2]) der Wahrscheinlichkeiten P für alle Kombinationen, wo dem A 1 bis 9 und dem B 1 bis 7 Partien fehlen; nachstehend ist ein Bruchstück der Tafel mitgeteilt, das sich mechanisch leicht fortsetzen läßt.[3])

		Dem B fehlen Partien			
		1	2	3	4
Dem A fehlen Partien	1	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$
	2	$\frac{1}{4}$	$\frac{4}{8}$	$\frac{11}{16}$	$\frac{26}{32}$
	3	$\frac{1}{8}$	$\frac{5}{16}$	$\frac{16}{32}$	$\frac{42}{64}$
	4	$\frac{1}{16}$	$\frac{6}{32}$	$\frac{22}{64}$	$\frac{64}{128}$
	5	$\frac{1}{32}$	$\frac{7}{64}$	$\frac{29}{128}$	$\frac{93}{256}$

Erst Moivre[4]) hat die Voraussetzung $p = q = \frac{1}{2}$ aufgegeben. Die allgemeine Lösung ist zuerst von Montmort[5]) in den beiden Formen (1) und (2) veröffentlicht worden. Die erste dieser Formeln beruht auf dem von Fermat angegebenen Gedanken.

Von einem andern Gedanken, der zu einer in späterer Zeit vielfach angewandten Methode der Lösung von Wahrscheinlichkeitsproblemen geführt hat, war Pascal ausgegangen. Man bezeichne die Wahrscheinlichkeit des Gewinnens für A, wenn ihm m Partien und seinem Gegner n Partien fehlen, mit $y_{m,n}$ und denke sich das Spiel noch um eine Partie fortgesetzt; A kann das Spiel gewinnen, wenn er diese Partie und auf der neuen Grundlage das Spiel gewinnt, wofür die Wahrscheinlichkeit $py_{m-1,n}$ ist; er kann es aber auch ge-

1) Ch. Huygens, De Ratiociniis in Ludo Aleae (1657). S. auch J. Bernoulli, Ars conjectandi (1713), pars prima.

2) Ars conjectandi, p. 16 (Ostwalds Klassiker Nr. 107, p. 18).

3) Das Fortschreiten der ersten Kolonne und Zeile ist unmittelbar ersichtlich; jede andere Zahl ist das arithmetische Mittel aus den ihr links und oben benachbarten.

4) Doctrine of chances (1718, 1738, 1756), p. 18.

5) Essai d'Analyse sur les Jeux de Hasards (1714), p. 232—248.

winnen, wenn er diese Partie verliert, das Spiel aber trotzdem auf der neuen Grundlage gewinnt, wofür die Wahrscheinlichkeit $qy_{m,n-1}$ ist; mithin besteht die Gleichung:

$$y_{mn} = py_{m-1,n} + qy_{m,n-1}. \tag{3}$$

Nun ist

$$y_{11} = p, \quad y_{21} = p^2, \quad y_{31} = p^3, \quad y_{41} = p^4, \cdots$$

weil $y_{0,\alpha} = 1$ und $y_{\beta,0} = 0$ ist, da A sicher gewinnt, wenn ihm keine Partie, und unmöglich gewinnen kann, wenn dem B keine Partie mehr fehlt; mit Hilfe derselben Bemerkung ergibt sich:

$$y_{12} = p + pq, \quad y_{13} = p + pq + pq^2, \quad y_{14} = p + pq + pq^2 + pq^3, \cdots$$

und weiter:

$$y_{22} = p^2(1 + 2q), \quad y_{32} = p^3(1 + 3q), \quad y_{42} = p^4(1 + 4q), \cdots$$
$$y_{33} = p^3(1 + 3q + 6q^2), \quad y_{43} = p^4(1 + 4q + 10q^2), \cdots$$

Auf diese Weise kann man zu jedem beliebigen $y_{mn} = P$ fortschreiten.

41. Beispiel XXV. *Über n Paare gleichartiger entgegengesetzter Ereignisse E_i, $\bar{E}_i$, deren Wahrscheinlichkeiten p_i, $q_i = 1 - p_i$ sind, wird eine Beobachtung oder ein Versuch angestellt. Wie groß ist die Wahrscheinlichkeit, daß genau k von den Ereignissen E, und wie groß die Wahrscheinlichkeit, daß mindestens k von den Ereignissen E eintreffen?*

Zur Erläuterung des Problems möge das folgende Beispiel dienen, dessen Lösung am Schlusse wird gegeben werden. Es liegen fünf äußerlich gleiche Urnen U_i mit folgendem Inhalte vor:

U_1	enthält	3	weiße,	2	schwarze	Kugeln,
U_2	„	2	„	1	„	Kugel,
U_3	„	4	„	1	„	„
U_4	„	1	„	2	„	Kugeln,
U_5	„	2	„	2	„	„ ;

man zieht aus jeder Urne eine Kugel heraus; wie groß ist die Wahrscheinlichkeit, daß sich unter den 5 gezogenen Kugeln genau 3, und daß sich darunter mindestens 3 weiße Kugeln befinden?

Um die erste Wahrscheinlichkeit — sie heiße $\mathfrak{w}(E^k)$ — zu finden, hätte man auf alle möglichen Arten Produkte von k Faktoren p_i und $n - k$ Faktoren q_i zu bilden und zu summieren; demnach kann P_k symbolisch durch die Formel

$$\mathfrak{w}(E^k) = \sum p_\alpha p_\beta \cdots p_\varkappa \cdot q_\lambda q_\mu \cdots q_\nu \tag{1}$$

dargestellt werden, wobei sich die Summenbildung auf alle Kombi-

nationen $\alpha, \beta, \cdots \varkappa$ der Elemente $1, 2, \cdots n$ zur k-ten Klasse bezieht; $\lambda, \mu, \cdots \nu$ sind die jeweilen in der Kombination nicht vorhandenen Elemente.

Man kann aber $\mathfrak{w}(E^k)$ auch durch die p_i allein darstellen, indem man schreibt:

$$\mathfrak{w}(E^k) = \sum p_\alpha p_\beta \cdots p_k(1 - p_\lambda)(1 - p_\mu) \cdots (1 - p_\nu); \tag{2}$$

entwickelt man das Produkt hinter dem Summenzeichen, so entstehen Produkte der p_i zu je k, je $k + 1, \cdots$ bis zu n Faktoren, und bei der darauf folgenden Summierung treten die Summen

$$S_k, \quad S_{k+1}, \cdots S_n$$

derartiger Produkte auf. Die Produkte von k Faktoren kommen nur je einmal vor, ihre Summe S_k erhält daher den Koeffizienten 1; jedes Produkt von mehr als k Faktoren tritt mehrfach auf, weil es auf mehrere Arten entstehen kann; so z. B. entsteht das Produkt $p_1 p_2 \cdots p_k p_{k+1}$ nicht bloß bei Entwicklung des Produktes

$$p_1 p_2 \cdots p_k(1 - p_{k+1})(1 - p_{k+2}) \cdots (1 - p_n),$$

sondern auch bei der Entwicklung jedes Produktes, dessen k erste Faktoren eine Kombination der k-ten Klasse aus den Elementen $p_1, p_2, \cdots p_k, p_{k+1}$ bilden; jede der Summen von S_{k+1} aufwärts erhält somit einen von 1 verschiedenen Koeffizienten, so daß in entwickelter Form

$$\mathfrak{w}(E^k) = S_k + A_1 S_{k+1} + A_2 S_{k+2} + \cdots + A_{n-k} S_n \tag{3}$$

sein wird.

Die Koeffizienten $A_1, A_2, \cdots$ hängen von den p_i nicht ab, bleiben also dieselben, wenn man alle p_i als gleich und gleich p voraussetzt; dann aber verwandelt sich die unentwickelte Summe (2) in

$$\binom{n}{k} p^k (1 - p)^{n-k}, \tag{4}$$

und die entwickelte Summe (3) geht, weil nun

$$S_k = \binom{n}{k} p^k, \quad S_{k+1} = \binom{n}{k+1} p^{k+1}, \cdots S_n = p^n$$

wird, über in

$$\binom{n}{k} p^k + A_1 \binom{n}{k+1} p^{k+1} + A_2 \binom{n}{k+2} p^{k+2} + \cdots + A_{n-k} p^n. \tag{5}$$

Durch Vergleichung der Koeffizienten gleicher Potenzen von p in (4) und (5) ergeben sich zur Bestimmung der A die Gleichungen:

$$-\binom{n}{k}\binom{n-k}{1} = A_1\binom{n}{k+1}, \quad \binom{n}{k}\binom{n-k}{2} = A_2\binom{n}{k-2},$$

$$-\binom{n}{k}\binom{n-k}{3} = A_3\binom{n}{k+3}, \quad \cdots (-1)^{n-k}\binom{n}{k}\binom{n-k}{n-k} = A_{n-k};$$

aus diesen berechnet sich

$$A_1 = -\binom{k+1}{1}, A_2 = \binom{k+2}{2}, A_3 = -\binom{k+3}{3}, \cdots A_{n-k} = (-1)^{n-k}\binom{n}{n-k}.$$

Somit ist schließlich

$$\mathfrak{w}(E^k) = S_k - \binom{k+1}{1}S_{k+1} + \binom{k+2}{2}S_{k+2} - \cdots + (-1)^{n-k}\binom{n}{n-k}S_n. \quad (6)$$

Entwickelt man den algebraischen Ausdruck

$$\frac{S^k}{(1+S)^{k+1}},$$

in welchem S das Zeichen für eine beliebige Größe vorstellt, so ergibt sich:

$$\frac{S^k}{(1+S)^{k+1}} = S^k(1+S)^{-(k+1)}$$

$$= S^k - \binom{k+1}{1}S^{k+1} + \binom{k+2}{2}S^{k+2} - \cdots + (-1)^{n-k}\binom{n}{n-k}S^n \cdots;$$

daher kann symbolisch

$$\mathfrak{w}(E^k) = \frac{S^k}{(1+S)^{k+1}} \quad (7)$$

gesetzt werden, wenn man festsetzt, daß die Entwicklung von $(1+S)^{-(k+1)}$ bei dem Gliede mit S^{n-k} abgebrochen und nachher jedes S^ω durch S_ω ersetzt wird.

Die Wahrscheinlichkeit, daß *mindestens* k Ereignisse der Gattung E eintreten, ist die vollständige Wahrscheinlichkeit dafür, daß genau k oder genau $k+1 \cdots$ oder genau n dieser Ereignisse eintreffen; bezeichnet man sie also mit $\mathfrak{W}(E^k)$, so ist

$$\begin{aligned}
\mathfrak{W}(E^k) &= \mathfrak{w}(E^k) + \mathfrak{w}(E^{k+1}) + \cdots + \mathfrak{w}(E^n) \\
&= S_k - \binom{k+1}{1}S_{k+1} + \binom{k+2}{2}S_{k+2} + \cdots + (-1)^{n-k}\binom{n}{n-k}S_n \\
&\quad + \binom{k+1}{0}S_{k+1} - \binom{k+2}{1}S_{k+2} + \cdots + (-1)^{n-k-1}\binom{n}{n-k-1}S_n \\
&\quad + \binom{k+2}{0}S_{k+2} - \cdots + (-1)^{n-k-2}\binom{n}{n-k-2}S_n \\
&\quad \cdots\cdots\cdots\cdots\cdots \\
&\quad + \binom{n}{0}S_n;
\end{aligned}$$

nun hat man aber

$$\binom{k+1}{0}-\binom{k+1}{1}=\binom{k}{0}-\binom{k+1}{1}=\binom{k}{0}-\binom{k}{0}-\binom{k}{1}=-\binom{k}{1}$$

$$\binom{k+2}{0}-\binom{k+2}{1}+\binom{k+2}{2}$$

$$=\binom{k+1}{0}-\binom{k+1}{0}-\binom{k+1}{1}+\binom{k+1}{1}+\binom{k+1}{2}=\binom{k+1}{2},\cdots$$

Demzufolge ist endgültig

$$\mathfrak{W}(E^k) = S_k - \binom{k}{1} S_{k+1} + \binom{k+1}{2} S_{k+2} - \cdots + (-1)^{n-k}\binom{n-1}{n-k} S_n. \quad (8)$$

Vergleicht man dies wie oben mit der Entwicklung:

$$\frac{S^k}{(1+S)^k} = S^k(1+S)^{-k}$$

$$= S^k - \binom{k}{1} S^{k+1} + \binom{k+1}{2} S^{k+2} - \cdots + (-1)^{n-k}\binom{n-1}{n-k} S^n + \cdots,$$

so ergibt sich die symbolische Darstellung:

$$\mathfrak{W}(E^k) = \frac{S^k}{(1+S)^k}. \quad (9)$$

In dem eingangs aufgestellten Urnenbeispiel ist $n = 5$ und $k = 3$,

$$\mathfrak{w}(E^3) = \frac{S^3}{(1+S)^4} = S_3 - 4S_4 + 10S_5,$$

$$\mathfrak{W}(E^3) = \frac{S^3}{(1+S)^3} = S_3 - 3S_4 + 6S_5;$$

weiter

$$S_3 = p_1p_2p_3 + p_1p_2p_4 + p_1p_2p_5 + p_1p_3p_4 + p_1p_3p_5 + p_1p_4p_5$$
$$+ p_2p_3p_4 + p_2p_3p_5 + p_2p_4p_5 + p_3p_4p_5 = \frac{829}{450},$$

$$S_4 = p_1p_2p_3p_4 + p_1p_2p_3p_5 + p_1p_2p_4p_5 + p_1p_3p_4p_5$$
$$+ p_2p_3p_4p_5 = \frac{226}{450},$$

$$S_5 = p_1p_2p_3p_4p_5 = \frac{24}{450},$$

weil $p_1 = \frac{3}{5}$, $p_2 = \frac{2}{3}$, $p_3 = \frac{4}{5}$, $p_4 = \frac{1}{3}$, $p_5 = \frac{1}{2}$ ist; demnach ist die Wahrscheinlichkeit, daß genau 3 weiße Kugeln gezogen werden:

$$\mathfrak{w}(E^3) = \frac{829}{450} - 4 \cdot \frac{226}{450} + 10 \cdot \frac{24}{450} = \frac{33}{90}$$

und die Wahrscheinlichkeit, daß mindestens 3 weiße Kugeln erscheinen:

$$\mathfrak{W}(E^3) = \frac{829}{450} - 3 \cdot \frac{226}{450} + 6 \cdot \frac{24}{450} = \frac{59}{90}.$$

42. Beispiel XXVI. *Es ist die Wahrscheinlichkeit zu bestimmen, mit n Würfeln eine bestimmte Summe zu werfen.*

In dem Polynom $\frac{1}{6}x^1 + \frac{1}{6}x^2 + \frac{1}{6}x^3 + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6$ sind die Fälle, welche sich mit *einem* Würfel zutragen können, und ihre Wahrscheinlichkeiten in eigentümlicher Weise zum Ausdruck gebracht: der Exponent des Hilfsbuchstaben x bedeutet die Anzahl Augen und der Koeffizient die Wahrscheinlichkeit, daß diese Anzahl sich einstelle.

Multipliziert man das Polynom mit sich selbst und ordnet das Produkt nach den Potenzen von x, so werden die Exponenten addiert, die Koeffizienten der miteinander multiplizierten Glieder multipliziert und die Koeffizienten gleicher Potenzen zu einer Summe vereinigt; die Rechnung geht also nach solchen Regeln vor sich, daß der Koeffizient von x^s in dem entwickelten Produkt unmittelbar die Wahrscheinlichkeit angibt, daß mit zwei Würfeln die Summe s getroffen werde. Multipliziert man das Produkt aufs neue mit dem Polynom, so gibt in der nach Potenzen von x geordneten Entwicklung der Koeffizient von x^s aus den gleichen Gründen die Wahrscheinlichkeit an, daß mit drei Würfeln die Summe s erzielt werde.

Allgemein wird demnach der Koeffizient von x^s in der Entwicklung von

$$\left(\frac{1}{6}x^1 + \frac{1}{6}x^2 + \frac{1}{6}x^3 + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6\right)^n$$

die Wahrscheinlichkeit vorstellen, daß mit n Würfeln die Summe s falle. Sondert man den Nenner 6^n ab, der die möglichen Fälle zählt, die sich bei einem Wurfe mit n Würfeln zutragen können, so bleibt

$$(x^1 + x^2 + x^3 + x^4 + x^5 + x^6)^n,$$

und nunmehr zählt der Koeffizient von x^s die der Summe s günstigen Fälle.

Dies ist die mathematische Grundlage des mechanischen Verfahrens, welches J. Bernoulli[1]) zur Bestimmung der den verschiedenen Summen günstigen Fälle angegeben hat; das Verfahren ist nichts anderes als die schematische Bildung der Koeffizienten, welche sich

1) Ars conjectandi, p. 24 (Ostwalds Klassiker Nr. 107, p. 27).

bei der sukzessiven Multiplikation des Polynoms $x^1 + x^2 + \cdots + x^6$ mit sich selbst ergeben; es ist aus der folgenden Tabelle, welche bis zu 3 Würfeln oder 3 Würfen mit einem Würfel reicht, zu ersehen. Dem Multiplikationsverfahren entsprechend ist die Koeffizientenreihe jedesmal sechsmal untereinander geschrieben mit gleichzeitiger Vorrückung um eine Stelle; darauf erfolgt kolonnenweise Addition.

Summen:

Würfel	I	1	2	3	4	5	6										
	II	2	3	4	5	6	7	8	9	10	11	12					
	III	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Anzahlen der Entstehungsarten	I	1	1	1	1	1	1										
			1	1	1	1	1	1									
				1	1	1	1	1	1								
					1	1	1	1	1	1							
						1	1	1	1	1	1						
							1	1	1	1	1	1					
	II	1	2	3	4	5	6	5	4	3	2	1					
			1	2	3	4	5	6	5	4	3	2	1				
				1	2	3	4	5	6	5	4	3	2	1			
					1	2	3	4	5	6	5	4	3	2	1		
						1	2	3	4	5	6	5	4	3	2	1	
							1	2	3	4	5	6	5	4	3	2	1
	III	1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1

Aus dieser Tabelle entnimmt man z. B., daß die Summe 5 mit einem Würfel auf 1, mit zwei Würfeln auf 4, mit drei Würfeln auf 6 Arten entstehen kann, so daß ihr in den drei Fällen die Wahrscheinlichkeit $\frac{1}{6}, \frac{4}{36}, \frac{6}{216}$ zukommt. (Vgl. hierzu Nr. 14 und 15.)

43. Beispiel XXVII. *In einer Urne befinden sich $n + 1$ mit den Nummern $0, 1, 2, \cdots n$ bezeichnete Kugeln; man führt nacheinander i Ziehungen aus und legt jedesmal die gezogene Kugel, nachdem man ihre Nummer notiert hat, zurück. Welches ist die Wahrscheinlichkeit, daß die Summe der erschienenen Nummern s sei?*

Diese Aufgabe, in der man unschwer eine Verallgemeinerung der vorigen Würfelaufgabe erkennt, wird in der Literatur als Moivres Problem bezeichnet, weil Moivre ihre allgemeine Lösung zuerst veröffentlicht hat.[1] Der Gedanke, auf welchen er die Lösung gegründet hat, ist der in der vorigen Nummer entwickelte: Die Zahl der Arten, auf welche die Summe s entstehen kann, ist durch den Koeffizienten von x^s in der Entwicklung von

1) A. de Moivre, De mensura sortis (1711), p. 364, und Miscellanea analytica (1730), p. 191 ff.

$$(x^0 + x^1 + x^2 + \cdots + x^n)^i \tag{1}$$

bezeichnet; und da $(n+1)^i$ die Anzahl der möglichen Fälle ist, so erhält man die verlangte Wahrscheinlichkeit durch Division jenes Koeffizienten mit $(n+1)^i$.

Nun läßt sich der Ausdruck (1), wenn man die in der Klammer eingeschlossene geometrische Reihe summiert, umformen in:

$$(1 - x^{n+1})^i (1-x)^{-i} = \left[1 - i x^{n+1} + \binom{i}{2} x^{2n+2} - \binom{i}{3} x^{3n+3} + \cdots\right]$$
$$\times \left[1 + ix + \binom{i+1}{2} x^2 + \binom{i+2}{3} x^3 + \cdots\right].$$

Bei der Ausführung dieser Multiplikation entstehen Glieder mit x^s dadurch, daß man

das 1. Glied des ersten Faktors multipliziert mit dem Gliede $\binom{i+s-1}{s} x^s$ des zweiten Faktors,

das 2. Glied des ersten Faktors multipl. m. d. Gliede $\binom{i+s-n-2}{s-n-1} x^{s-n-1}$ des zweiten Faktors,

das 3. Glied des ersten Faktors multipliziert mit dem Gliede $\binom{i+s-2n-3}{s-2n-2} x^{s-2n-2}$ des zweiten Faktors,

. . . .

mithin ist der Koeffizient von x^s gleich

$$\binom{i+s-1}{s} - \binom{i}{1}\binom{i+s-n-2}{s-n-1} + \binom{i}{2}\binom{i+s-2n-3}{s-2n-2} - \cdots; \tag{2}$$

darin ist

$$\binom{i+s-1}{s} = \frac{(i+s-1)!}{s!\,(i-1)!} = \frac{(s+1)(s+2)\cdots(s+i-1)}{1\cdot 2\cdots(i-1)}$$

$$\binom{i+s-n-2}{s-n-1} = \frac{(i+s-n-2)!}{(s-n-1)!\,(i-1)!} = \frac{(s-n)(s-n+1)\cdots(s-n+i-2)}{1\cdot 2\cdots(i-1)}$$

$$\binom{i+s-2n-3}{s-2n-2} = \frac{(i+s-2n-3)!}{(s-2n-2)!\,(i-1)!} = \frac{(s-2n-1)(s-2n)\cdots(s-2n+i-3)}{1\cdot 2\cdots(i-1)}$$

. . . .

so daß für (2) auch

$$\frac{(s+1)(s+2)\cdots(s+i-1)}{1\cdot 2\cdots(i-1)} - \binom{i}{1}\frac{(s-n)(s-n+1)\cdots(s-n+i-2)}{1\cdot 2\cdots(i-1)}$$
$$+ \binom{i}{2}\frac{(s-2n-1)(s-2n)\cdots(s-2n+i-3)}{1\cdot 2\cdots(i-1)} - \cdots$$

geschrieben werden kann; die Reihe bricht ab, sobald im Zähler negative Faktoren auftreten sollten.

Die verlangte Wahrscheinlichkeit ist demnach

$$P_s = \frac{1}{(n+1)^i}\left\{\frac{(s+1)(s+2)\cdots(s+i-1)}{1\cdot 2\cdots(i-1)} - \binom{i}{1}\frac{(s-n)(s-n+1)\cdots(s-n+i-2)}{1\cdot 2\cdots(i-1)}\right.$$
$$\left. + \binom{i}{2}\frac{(s-2n-1)(s-2n)\cdots(s-2n+i-3)}{1\cdot 2\cdots(i-1)} - \cdots\right\}. \quad (3)$$

Wären z. B. in einer Urne 11 mit $0, 1, 2, \cdots 10$ beschriebene Zettel, und würde man dreimal einen Zettel ziehen, so bestünde für die Summe 25 Wahrscheinlichkeit

$$P_{25} = \frac{1}{11^3}\left\{\frac{26\cdot 27}{1\cdot 2} - \frac{3}{1}\frac{15\cdot 16}{1\cdot 2} + \frac{3\cdot 2}{1\cdot 2}\frac{4\cdot 5}{1\cdot 2}\right\} = \frac{21}{1331}.$$ [1]

Das Moivresche Problem läßt eine wichtige Verallgemeinerung in folgender Richtung zu. Während nach seiner bisherigen Fassung jeder Zug oder Versuch nur einen Wert aus der ganzzahligen Reihe von 0 bis n ergeben konnte, möge nun jeder reelle Zahlenwert aus einem bezeichneten Intervall $(0, N)$ zulässig und jeder an sich gleich wahrscheinlich sein; die Summe S der Ergebnisse von i Versuchen kann dann jede reelle Zahl aus dem Intervall $(0, iN)$ sein. Um

1) Um auf den Fall überzugehen, daß die Bezeichnung der Kugeln oder Zettel statt mit 0 mit 1 beginnt und daher bis $n+1$ fortschreitet, braucht man sich nur auf jeder Kugel die Nummer um 1 erhöht zu denken; dadurch erhöht sich die Summe aus i Ziehungen um i und wird $s+i$ statt s; nennt man sie nunmehr s', so ist $s = s' - i$ zu setzen, und Formel (3), die nun auch die Wahrscheinlicheit dieser Summe bestimmt, verwandelt sich in:

$$P_{s'} = \frac{1}{(n+1)^i}\left\{\frac{(s'-i+1)(s'-i+2)\cdots(s'-1)}{1\cdot 2\cdots(i-1)}\right.$$
$$- \binom{i}{1}\frac{(s'-i-n)(s'-i-n+1)\cdots(s'-n-2)}{1\cdot 2\cdots(i-1)}$$
$$\left. + \binom{i}{2}\frac{(s'-i-2n-1)(s'-i-2n)\cdots(s'-2n-3)}{1\cdot 2\cdots(i-1)} - \cdots\right\}. \quad (3^*)$$

Diese Formel stellt nun auch die allgemeine Lösung des Würfelproblems in Nr. 42 dar. Um beispielsweise die Wahrscheinlichkeit zu finden, mit 3 Würfeln die Summe 8 zu werfen, hat man $n = 5$, $i = 3$, $s' = 8$, und man erhält:

$$P_8 = \frac{1}{6^3}\frac{6\cdot 7}{1\cdot 2} = \frac{21}{216}$$

in Übereinstimmung mit der dort mitgeteilten Tafel; und um die Wahrscheinlichkeit zu berechnen, mit 6 Würfeln die Summe 15 zu treffen, für welche die Tafel nicht ausreicht, hat man $n = 5$, $i = 6$, $s' = 15$ zu setzen und findet:

$$P_{15} = \frac{1}{6^6}\left\{\frac{10\cdot 11\cdot 12\cdot 13\cdot 14}{1\cdot 2\cdot 3\cdot 4\cdot 5} - \frac{6}{1}\frac{4\cdot 5\cdot 6\cdot 7\cdot 8}{1\cdot 2\cdot 3\cdot 4\cdot 5}\right\} = \frac{1666}{46656} = \frac{833}{23328}.$$

nun die Wahrscheinlichkeit eines bestimmten Wertes S dieser Summe zu gewinnen, ändern wir zunächst die Bedingungen der Aufgabe dahin ab, daß die Kugeln in der Urne statt mit $0, 1, 2, \cdots n$ mit 0θ, 1θ, 2θ, $\cdots n\theta$ bezeichnet seien; P_s ist dann die Wahrscheinlichkeit, daß sich in i Ziehungen die Summe $s\theta$ ergibt, und ihr Ausdruck kann, damit er diese Summe enthalte, umgeformt werden in:

$$P_{s\theta} = \frac{\theta}{1 \cdot 2 \cdots (i-1)(n\theta+\theta)} \left\{ \frac{(s\theta+\theta)(s\theta+2\theta)\cdots(s\theta+i\theta-\theta)}{(n\theta+\theta)^{i-1}} \right.$$

$$- \binom{i}{1} \frac{(s\theta-n\theta)(s\theta-n\theta+\theta)\cdots(s\theta-n\theta+i\theta-2\theta)}{(n\theta+\theta)^{i-1}}$$

$$\left. + \binom{i}{2} \frac{(s\theta-2n\theta-n\theta)(s\theta-2n\theta)\cdots(s\theta-2n\theta+i\theta-3\theta)}{(n\theta+\theta)^{i-1}} - \cdots \right\}.$$

Läßt man nun n und s ins Unendliche wachsen, θ gleichzeitig gegen Null abnehmen, jedoch so, daß

$$\lim n\theta = N, \quad \lim s\theta = S$$

wird, so bedeutet N die höchste unter den möglichen Zahlen, S die Summe der aus i Ziehungen stammenden Zahlen und

$$\theta = dS$$

die geringste Änderung, welcher diese Summe fähig ist. Das Resultat dieses Grenzüberganges:

$$P_s = \frac{dS}{1 \cdot 2 \cdots (i-1) N} \left\{ \left(\frac{S}{N}\right)^{i-1} - \binom{i}{1}\left(\frac{S}{N}-1\right)^{i-1} + \binom{i}{2}\left(\frac{S}{N}-2\right)^{i-1} - \cdots \right\}$$

ist zu deuten als Wahrscheinlichkeit einer zwischen S und $S+\theta = S+dS$ liegenden Summe; das Integral hiervon, genommen zwischen 0 und S, bedeutet dann die totale Wahrscheinlichkeit, daß die Summe in eines der Intervalle $(0, \theta)$ oder $(\theta, 2\theta)$, $\cdots$ oder $(S-\theta, S)$, d. h. die Wahrscheinlichkeit, daß sie zwischen 0 und S falle; diese ist sonach

$$P = \frac{1}{1 \cdot 2 \cdots i} \left\{ \left(\frac{S}{N}\right)^{i} - \binom{i}{1}\left(\frac{S}{N}-1\right)^{i} + \binom{i}{2}\left(\frac{S}{N}-2\right)^{i} - \cdots \right\}. \quad (4)$$

Die Entwicklung ist so weit zu führen, als die Differenzen positiv bleiben.

Auf diese Formel hat Laplace[1]) die Lösung einer Frage gegründet, die mit einer 1732 und 1734 von der Pariser Akademie gestellten Preisaufgabe zusammenhängt. Die Summe der Neigungen

1) Théorie analyt., Art. 13.

der 10 Planetenbahnen gegen die Ekliptik war zu Beginn des vorigen Jahrhunderts mit 91,4187 Zentesimalgraden bestimmt; diese Summe ist bemerkenswerterweise kleiner als die Neigung einer einzelnen Bahn sein *könnte*. Man kann nun die Frage aufstellen: Wären jene Bahnen ein Werk des Zufalls und bestünde kein Grund, warum ein Wert der Neigung zwischen 0° und 100° leichter möglich sein sollte als ein anderer, welche Wahrscheinlichkeit bestünde dann, daß die Summe aller 10 Neigungen unter 92° liegt? Die Formel (4) gibt hierauf als Antwort:

$$P = \frac{1}{1 \cdot 2 \cdots 10}\left(\frac{92}{100}\right)^{10} = 0{,}00000012.$$

Besteht hingegen ein innerer Grund, der gerade diese Neigungen notwendig macht, dann kommt der beobachteten Summe die Wahrscheinlichkeit 1 zu, die sich zur vorigen verhält etwa wie 25 Millionen zu 3; dieses Resultat, so bemerkt Laplace, weist mit sehr großer Wahrscheinlichkeit auf das Vorhandensein einer ursprünglichen Ursache hin, welche die Bewegung der Planeten so bestimmt hat, daß sich ihre Bahnen nahe der Ekliptik angeordnet haben. Übrigens hat Laplace zur Verstärkung dieser Vermutung, wenn hier von Vermutung überhaupt gesprochen werden kann, auch auf den gleichen Umlaufsinn der Planeten hingewiesen und die zusammengesetzte Wahrscheinlichkeit gerechnet, daß nicht nur die Summe der Neigungen unter der obigen Grenze liege, sondern auch der Umlaufsinn derselbe sei, letzteres wieder unter der Annahme, daß beide Bewegungsrichtungen gleich möglich seien.

44. Beispiel XXVIII. *Bei dem Bouillotespiel werden aus einem Spiel von 32 Blättern die Siebner, Zehner und Buben ausgeschieden und von den übrigen Blättern je 3 an die vier Spieler verteilt; das 13. Blatt wird umgewendet. Ein Spieler hat einen Brelan, wenn er drei gleichartige Blätter (As, Könige o. dgl.) erhält, und er hat einen Brelan carré, wenn die Blätter auch noch von der Art des umgewendeten sind. Es ist die Wahrscheinlichkeit zu bestimmen, 1) daß i bezeichnete Spieler einen Brelan haben; 2) daß i bezeichnete Spieler und nur diese Brelan haben; 3) daß bei einer bestimmten Reihenfolge der Spieler der i-te einen Brelan hat, ohne daß die vorausgehenden einen solchen besitzen.*

1) Daß ein erster Spieler — und dies kann jeder von den vier Spielern sein, weil die Art der Verteilung willkürlich ist — einen Brelan erhält, hat die Wahrscheinlichkeit

$$p_1 = \frac{20}{20} \cdot \frac{3}{19} \cdot \frac{2}{18} = 0{,}0175438;$$

denn welches Blatt er als erstes bekommt, die zwei andern Blätter

müssen von derselben Art sein; und solcher gibt es nach dem ersten Blatt noch 3 unter 19, nach dem zweiten noch 2 unter 18.

Soll ein zweiter Spieler einen Brelan erhalten, so darf er von den 17 noch übrigen Blättern als erstes dasjenige nicht erhalten, welches zu dem ersten Brelan als viertes gehört; da also ein Blatt ausgeschlossen ist, so beträgt die Wahrscheinlichkeit, daß außer dem ersten noch ein zweiter Spieler Brelan erhält,

$$p_2 = p_1 \frac{16}{17} \cdot \frac{3}{16} \cdot \frac{2}{15} = \frac{2}{85} p_1 = 0{,}0004128.$$

Durch ähnliche Schlüsse ergeben sich die Wahrscheinlichkeiten p_3, p_4 dafür, daß drei und daß alle vier Spieler einen Brelan erhalten, nämlich:

$$p_3 = p_2 \frac{12}{14} \cdot \frac{3}{13} \cdot \frac{2}{12} = \frac{3}{91} p_2 = 0{,}0000136,$$

$$p_4 = p_3 \frac{8}{11} \cdot \frac{3}{10} \cdot \frac{2}{9} = \frac{8}{165} p_3 = 0{,}0000007.$$

2) Die Wahrscheinlichkeit p_1, daß ein bezeichneter Spieler Brelan hat, besteht aus der Wahrscheinlichkeit w_1, daß er allein ihn hat, ferner aus der Wahrscheinlichkeit $3w_2$ [1]), daß er und ein zweiter und nur diese zwei ihn haben, dann aus der Wahrscheinlichkeit $3w_3$ [1]), daß er und noch zwei andere und nur diese drei ihn haben, endlich aus der Wahrscheinlichkeit $w_4 = p_4$, daß alle vier ihn haben; denn dies sind die voneinander unabhängigen Arten, auf welche der bezeichnete Spieler zu einem Brelan kommen kann; mithin ist

$$p_1 = w_1 + 3w_2 + 3w_3 + w_4;$$

ebenso erkennt man, daß

$$p_2 = w_2 + 2w_3 + w_4,$$

$$p_3 = w_3 + w_4;$$

endlich ist, wie schon bemerkt worden,

$$p_4 = w_4.$$

Hieraus ergeben sich die Wahrscheinlichkeiten, daß ein, zwei, drei Spieler und nur diese einen Brelan erhalten:

$$w_1 = p_1 - 3p_2 + 3p_3 - p_4 = 0{,}0163455,$$

$$w_2 = p_2 - 2p_3 + p_4 = 0{,}0003863,$$

$$w_3 = p_3 - p_4 = 0{,}0000129.$$

---

1) Der Koeffizient 3 rührt daher, daß der bezeichnete Spieler mit drei andern, beziehungsweise mit drei Paaren aus den übrigen Spielern sich verbinden kann.

3) Die Wahrscheinlichkeit P_4, daß der vierte Spieler einen Brelan erhält, ohne daß die vorangehenden einen solchen haben, fällt zusammen mit der Wahrscheinlichkeit w_1, daß ein Spieler allein einen Brelan macht. Die Wahrscheinlichkeit P_3, daß der dritte Spieler in der Reihe Brelan macht, ohne daß die vorangehenden einen solchen erhalten, setzt sich als Summe zusammen aus der Wahrscheinlichkeit w_1, daß er ihn allein macht, und aus der Wahrscheinlichkeit w_2, daß er ihn mit dem nachfolgenden Spieler zugleich erhält. Bei dem zweiten Spieler ist darauf zu achten, daß er Brelan entweder allein oder zusammen mit einem der beiden folgenden oder mit beiden zugleich haben kann. Der erste Spieler endlich kann entweder allein oder mit einem oder mit einem Paar der drei folgenden oder mit allen zugleich Brelan machen. Hiernach ist

$$P_4 = w_1 = 0{,}0163455$$
$$P_3 = w_1 + w_2 = 0{,}0167318$$
$$P_2 = w_1 + 2w_2 + w_3 = 0{,}0171310$$
$$P_1 = w_1 + 3w_2 + 3w_3 + w_4 = 0{,}0175438.$$

Die Wahrscheinlichkeit Π, daß mindestens ein Brelan zustande kommt, ist die totale Wahrscheinlichkeit dafür, daß ein oder daß zwei oder drei oder alle vier Spieler Brelan erhalten, was beziehungsweise wieder auf 4, 6, 4 Arten und 1 Art geschehen kann; mithin ist

$$\Pi = 4w_1 + 6w_2 + 4w_3 + w_4 = 0{,}0677521;$$

daraus folgt die Wahrscheinlichkeit, daß kein Brelan gemacht werde, mit 0,9322479.

45. Beispiel XXIX. *Im „Trente et quarante" genannten Spiele gibt es nur einen Fall, welcher dem Bankhalter günstig ist; er besteht darin, daß zweimal nacheinander die Summe 31 zustande kommt. Es ist die Wahrscheinlichkeit dieses Falles zu bestimmen.*

Das Spiel wird mit der Vereinigung mehrerer, durcheinander gemischter vollständiger Spiele von je 52 Karten ausgeführt, etwa mit 8 Spielen oder 416 Karten. Die Figuren gelten 10, die übrigen Blätter entsprechend den auf ihnen verzeichneten Points.

Die Karten werden einzeln aufgeschlagen so lange, bis eine 30 übersteigende, höchstens also die Summe 40 entsteht. Derselbe Vorgang wird ein zweites Mal wiederholt. Der Spieler wettet auf die erste oder die zweite Summe und gewinnt, wenn sie die größere von beiden ist.

Kommen zwei gleiche Summen heraus, so ist das Spiel ungültig. Sind es aber zwei 31, so zieht der Bankhalter die Hälfte des Einsatzes ein. Um die Wahrscheinlichkeit dieses Falles handelt es sich,

Die Lösung der Aufgabe würde sich außerordentlich verwickelt gestalten, wollte man auf die Veränderung Rücksicht nehmen, welche die Wahrscheinlichkeit einer Karte bestimmter Art während des Spieles erleiden kann je nach den vor ihr schon aufgeschlagenen Karten. Da aber der Einfluß dieser Veränderlichkeit auf das Endresultat vermöge der großen Gesamtzahl der Blätter und auch der großen Zahl der Blätter einer bestimmten Art nur unerheblich sein kann, so darf man sich auf den Standpunkt stellen, es bleibe die Wahrscheinlichkeit für eine Karte bestimmter Art, z. B. einen König, durch die ganze Dauer des Spiels dieselbe, $\frac{1}{13}$.

Es bezeichne p_s die Wahrscheinlichkeit, die Summe s zu machen.

Die Summe 1 ergibt sich nur, wenn ein As gezogen wird, also ist

$$p_1 = \frac{1}{13}.$$

Die Summe 2 kann entstehen, entweder indem ein Zweier oder zweimal As gezogen wird, also ist

$$p_2 = \frac{1}{13} + \frac{1}{13^2} = \frac{1}{13}\left(1 + \frac{1}{13}\right).$$

Für die Summe 3 gibt es drei Entstehungsweisen: entweder ein Dreier, oder ein As und ein Zweier, oder nach Erzielung der Summe 2 ein As; daher ist

$$p_3 = \frac{1}{13} + \frac{1}{13^2} + \frac{1}{13}\left(1 + \frac{1}{13}\right) \cdot \frac{1}{13} = \frac{1}{13}\left(1 + \frac{1}{13}\right)^2.$$

Die Summe 4 kann auf vier Arten entstehen; entweder wird ein Vierer gezogen, oder es folgt der Summe 1 ein Dreier, der Summe 2 ein Zweier oder der Summe 3 ein As; folglich ist

$$p_4 = \frac{1}{13} + \frac{1}{13}p_1 + \frac{1}{13}p_2 + \frac{1}{13}p_3 = \frac{1}{13}\left(1 + \frac{1}{13}\right)^3.$$

Dies geht so fort bis zur Summe 9; entsprechend ihren neun Entstehungsarten ist

$$p_9 = \frac{1}{13} + \frac{1}{13}p_1 + \frac{1}{13}p_2 + \cdots + \frac{1}{13}p_8 = \frac{1}{13}\left(1 + \frac{1}{13}\right)^8.$$

Von der Summe 10 ab ändert sich die Sachlage, weil viererlei verschiedene Karten die Summe um 10 erhöhen können: Zehner, Buben, Königinnen und Könige.

Die Summe 10 selbst kann entstehen, indem man einen Zehner zieht oder indem der Summe 1 ein Neuner, der Summe 2 ein Achter, ⋯ endlich der Summe 9 ein As folgt; demnach ist

$$p_{10} = \frac{4}{13} + \frac{1}{13}(p_1 + p_2 + \cdots + p_9).$$

Für jede 10 übersteigende Summe gibt es folgende Entstehungsarten: entweder verbindet sich die Summe $s - 10$ mit einem Zehner oder es folgt der Summe $s - 9$ ein Neuner, der Summe $s - 8$ ein Achter $\cdots$ oder endlich der Summe $s - 1$ ein As; hiernach ist

$$p_s = \frac{4}{13} p_{s-10} + \frac{1}{13}(p_{s-9} + p_{s-8} + \cdots + p_{s-1}), \qquad (s > 10).$$

Durch Benutzung dieser Formeln findet man sukzessive:

$p_1 = 0{,}076923$	$p_{11} = 0{,}119980$	$p_{21} = 0{,}139847$
$p_2 = 0{,}082840$	$p_{12} = 0{,}124657$	$p_{22} = 0{,}142454$
$p_3 = 0{,}089213$	$p_{13} = 0{,}129344$	$p_{23} = 0{,}144905$
$p_4 = 0{,}096075$	$p_{14} = 0{,}134015$	$p_{24} = 0{,}147180$
$p_5 = 0{,}103465$	$p_{15} = 0{,}138639$	$p_{25} = 0{,}149259$
$p_6 = 0{,}111424$	$p_{16} = 0{,}143181$	$p_{26} = 0{,}151124$
$p_7 = 0{,}119995$	$p_{17} = 0{,}147603$	$p_{27} = 0{,}152756$
$p_8 = 0{,}129226$	$p_{18} = 0{,}151856$	$p_{28} = 0{,}154134$
$p_9 = 0{,}139166$	$p_{19} = 0{,}155890$	$p_{29} = 0{,}155240$
$p_{10} = 0{,}380641$	$p_{20} = 0{,}212902$	$p_{30} = 0{,}168347$
	$p_{31} = 0{,}148060.$	

Demnach ist die Wahrscheinlichkeit, zweimal nacheinander 31 zu machen:

$$p_{31}^2 = 0{,}0219220.$$

Poisson[1]) führte die Rechnung mit Berücksichtigung der Veränderlichkeit der Wahrscheinlichkeiten durch und fand als angenäherten Wert 0,021967; dabei war vorausgesetzt, daß acht Kartenspiele zusammengelegt seien. Man erkennt hieraus, wie gering der Einfluß der Variationen in der Anzahl der möglichen und günstigen Fälle während des Spieles ist.

46. Beispiel XXX. *Es liegen 5 äußerlich gleiche Urnen vor; drei davon (A) enthalten je 3 weiße (w) und 2 schwarze (s) Kugeln, die zwei andern (B) je eine weiße und 3 schwarze Kugeln. Man zieht aus einer der Urnen eine Kugel heraus, legt sie, ohne sie besehen zu haben, in eine andere Urne und zieht aus dieser, nachdem man ihren Inhalt durcheinander gemengt, eine Kugel. Welches ist die Wahrscheinlichkeit, daß sie weiß sei?*

---

1) Sur le jeu de trente et quarante. Annal. de Gergonne, XVI.

Das erwartete Ereignis kann auf 8 einander ausschließende Arten eintreffen, welche schematisch durch folgende Symbole dargestellt sind:

$$AwA, \quad AsA, \quad AwB, \quad AsB,$$
$$BwA, \quad BsA, \quad BwB, \quad BsB;$$

AwA bedeutet, daß die erste Kugel aus einer der Urnen A genommen wird, daß sie weiß sei und wieder in eine der Urnen A gelegt wird, aus der dann die zweite Ziehung gemacht wird; ähnlich die anderen Kombinationen.

Jede der acht Entstehungsarten ist wieder als ein zusammengesetztes Ereignis aufzufassen, bestehend aus vier einfachen Ereignissen; so erfordert die erste Entstehungsart, daß man in eine Urne der Gattung A greife $\left(\text{Wahrscheinlichkeit } \frac{3}{5}\right)$, daß man aus ihr eine weiße Kugel nehme $\left(\text{W. } \frac{3}{5}\right)$, daß man diese in eine andere der Urnen A lege $\left(\text{W. } \frac{1}{2}\right)$ und daß man aus dieser, die dann 4 weiße und 2 schwarze Kugeln enthält, eine weiße Kugel ziehe $\left(\text{W. } \frac{2}{3}\right)$.

Demnach ist die vollständige Wahrscheinlichkeit des erwarteten Ereignisses:

$$\frac{3}{5}\,\frac{3}{5}\,\frac{1}{2}\,\frac{2}{3}+\frac{3}{5}\,\frac{2}{5}\,\frac{1}{2}\,\frac{1}{2}+\frac{3}{5}\,\frac{3}{5}\,\frac{1}{2}\,\frac{2}{5}+\frac{3}{5}\,\frac{2}{5}\,\frac{1}{2}\,\frac{4}{5}$$
$$+\frac{2}{5}\,\frac{1}{4}\,\frac{3}{4}\,\frac{2}{3}+\frac{2}{5}\,\frac{3}{4}\,\frac{3}{4}\,\frac{1}{2}+\frac{2}{5}\,\frac{1}{4}\,\frac{1}{4}\,\frac{2}{5}+\frac{2}{5}\,\frac{3}{4}\,\frac{1}{4}\,\frac{1}{5}=\frac{927}{2000}.$$

§ 4. Geometrische Wahrscheinlichkeit.

47. Erweiterung der Wahrscheinlichkeitsdefinition. Es ist schon in Nr. 7 bemerkt worden, daß die dort gegebene Wahrscheinlichkeitsdefinition den Wahrscheinlichkeitsbegriff nicht erschöpft und daher nicht zur Erledigung aller Wahrscheinlichkeitsfragen ausreicht. Ja, es ist nur ein sehr kleiner Teil dieser Fragen, auf den jene Definition, die deutlich die Spur ihrer Entstehung aus Glücksspielen an sich trägt, zur Anwendung gebracht werden kann; Glücksspiele und ihnen analoge Einrichtungen bilden denn auch ihr vornehmliches Anwendungsgebiet.

Sobald man über dieses hinausgreift, stellt sich die Notwendigkeit einer Modifikation oder einer Erweiterung der Definition ein. Jede erweiterte Definition muß aber, soll sie dem Begriff der Wahrscheinlichkeit gerecht werden, die ursprüngliche Definition mit umfassen. Von der Gestaltung der Definition hängt dann die mathematische Bestimmung der Wahrscheinlichkeit ab.

Das tägliche Leben veranlaßt uns sehr häufig wenn auch nicht zu Wahrscheinlichkeitsbestimmungen, so doch zu Wahrscheinlichkeitsschätzungen. Wir urteilen über das Eintreffen oder Nichteintreffen eines Ereignisses auf Grund allgemeiner Erwägungen, bei denen wir uns durch erworbene Erfahrungen leiten, vielfach aber auch unseren Seelenzustand mit zu Worte kommen lassen; das Ergebnis wird nur ausnahmsweise eine numerische Wahrscheinlichkeit, sondern in der Regel lediglich eine allgemeine Aussage über die Größenbeziehung der Wahrscheinlichkeiten für das Eintreffen und das Nichteintreffen sein. Der subjektive Charakter eines so gewonnenen Resultates schließt einen erkenntnistheoretischen Wert desselben aus; unser Handeln ist aber durch derlei abschätzende Urteile vielfach beeinflußt. Wenn z. B. jemand vor der Ausführung eines vom morgigen Wetter abhängigen Planes steht, so wird er auf Grund seiner allgemeinen Kenntnisse über den Witterungsverlauf eine Schätzung der Aussichten auf gutes oder schlechtes Wetter vornehmen; der lebhafte Wunsch, den Plan ausführen zu können, die Furcht, ihn vereitelt zu sehen, werden auf die Schätzung Einfluß üben, ihr Ergebnis kann zum Aufgeben des Planes führen; ein Erkenntniswert kommt dem Resultat aber nicht zu.

Solche Wahrscheinlichkeitsschätzungen subjektiven Charakters fallen außerhalb des Rahmens einer Theorie.[1])

Besitzt die einem Problem zugrunde liegende Materie eine solche Struktur, daß sie die Aufstellung von möglichen Fällen, die Abtrennung von günstigen Fällen und die Angabe der relativen Möglichkeit der einzelnen Fälle gestattet, so ist eine apriorische Wahrscheinlichkeits*bestimmung* im eigentlichen Sinne möglich, und ihr Erkenntniswert um so höher, je weniger willkürliche Momente in den angeführten Prozessen zur Geltung kommen; zur Gänze werden solche Momente, wo es sich um physische Vorgänge und nicht lediglich um kombinatorische Fragen handelt, wie sie etwa die Zahlentheorie bietet, sich nicht ausschließen lassen.

Die Definition wird sich im vorliegenden Falle wie folgt gestalten:

Können aus der Verwirklichung der allgemeinen Bedingungen die Fälle

$$A_i (i = 1, 2, \cdots m)$$

hervorgehen; besitzen unter diesen die Fälle

$$A_x (x = 1, 2, \cdots g)$$

---

1) R. Lämmel, Untersuchungen über die Ermittlung von Wahrscheinlichkeiten (Inaug.-Dissert., 1904), p. 2, bezeichnet derlei Schätzungen als *intuitive Methode* der Wahrscheinlichkeitsbestimmung. Das Wort Methode scheint mir nicht am Platze zu sein, weil es sich nicht um ein geregeltes Verfahren handelt.

diejenige Eigenschaft, welche ihren Eintritt zum Ereignis E stempelt; ist endlich durch die Werte der Funktion

$$\varphi(i)(i = 1, 2, \cdots m)$$

die relative Möglichkeit der Fälle gekennzeichnet, so ist

$$\mathfrak{W}(E) = \frac{\sum_1^g \varphi(i)}{\sum_1^m \varphi(i)}. \tag{1}$$

In der Tat führt dieser Ansatz wieder zur ursprünglichen Definition, wenn $\varphi(i)$ *konstant*, die möglichen Fälle also gleich möglich sind, indem unter dieser Voraussetzung

$$\mathfrak{W}(E) = \frac{g}{m} \tag{2}$$

wird.[1])

48. Unendlich viele Fälle. Es handelt sich nun darum, ob der Wahrscheinlichkeitsbegriff eine natürliche Erklärung und die Definition eine vernunftgemäße Erweiterung zuläßt, wenn es sich um Ereignisse handelt, die unendlich vieler Erscheinungsformen fähig sind.

Stellen wir uns zunächst auf den Standpunkt, daß wir den einzelnen Erscheinungsformen Gleichberechtigung zusprechen, gleichgültig, ob aus zwingenden Gründen oder aus mangelndem Grunde einer gegenteiligen Annahme. Sagt man schlechthin, m und g *seien* unendlich, so ist mit dem Bruche $\frac{g}{m}$ nichts anzufangen. Beschränken wir uns aber zunächst auf eine *endliche* Anzahl μ der möglichen Fälle und die in ihr auftretende Anzahl γ günstiger Fälle, so hat der Bruch $\frac{\gamma}{\mu}$ einen bestimmten Wert; hört er nicht auf bestimmt zu sein, wenn μ, endlich bleibend, immer größer und größer wird, so mißt der Bruch die Wahrscheinlichkeit des Ereignisses nach der gewöhnlichen Definition, und konvergiert er dabei gegen eine bestimmte Grenze p, so entspricht es dem gesunden Menschenverstande, diese Grenze als die Wahrscheinlichkeit des betrachteten Ereignisses in dem Sinne zu

---

1) R. Lämmel, l. c., p. 15, bezeichnet den in dieser Definition normierten Vorgang als *Hypothesenprozeß* und erblickt in ihm den allgemeinen Typus der apriorischen Wahrscheinlichkeitsbestimmung. Er schematisiert ihn wie folgt:

Hypothese H_1: Aufstellung der möglichen Fälle.
Hypothese H_2: Aufstellung der günstigen Fälle.
Hypothese H_3: Angabe des *elementaren Gewichts* der einzelnen Fälle oder der *Valenzfunktion* $\varphi(i)$.

Als Hypothese H_4 tritt dann die Bildungsweise von $\mathfrak{W}(E)$ hinzu.

deuten, daß, wenn man μ Erscheinungsformen als möglich annimmt, wobei μ *immer endlich* bleibt, die Wahrscheinlichkeit durch eine Zahl gemessen wird, die wohl *nicht* p ist, aber p beliebig nahe gebracht werden kann. In Zeichen:

$$\mathfrak{W}(E) = \lim_{\mu = \infty} \frac{\gamma}{\mu} = p. \tag{3}$$

Dazu ist aber die wichtige Bemerkung zu machen, daß die Auswahl der als möglich erachteten Erscheinungsformen nicht beliebig getroffen werden darf, sondern so erfolgen muß, daß sie der Natur des Problems, insbesondere der Forderung der gleichen Möglichkeit entspricht.

Dieser Gedankengang soll an einer der ersten in geometrischem Gewande gestellten Fragen beleuchtet werden. Sie ist auf die Wahrscheinlichkeit gerichtet, daß eine Münze, die auf einen mit gleichen Figuren bedeckten horizontalen Boden geworfen wird, über die Grenzlinien zu liegen komme.[1]

Der Mittelpunkt der Münze kann auf jeden Punkt der Ebene auftreffen, es gibt also unendlich viele mögliche, aber auch unendlich viele günstige Fälle. Denkt man sich die Ebene mit einem quadratischen (oder in anderer Weise regulären) Gitter von Punkten überzogen und nimmt an, nur diese könnten getroffen werden, so ist durch diese Anordnung der Vorstellung, daß gleich große Teile der Ebene auch gleichberechtigte Gebiete für die Lage des Mittelpunktes der Münze bilden, um so genauer entsprochen, je kleiner die Entfernung x der Punkte ist, Figur 2. Ist nun das Rechteck $ABCD$ eine der Figuren, $A'B'C'D'$ ein zweites Rechteck, dessen Seiten von denen des früheren um den Radius der Münze entfernt sind, so hat man die Punkte zu zählen, die einerseits in $ABCD$ und andererseits in dem von beiden Rechtecken gebildeten Rahmen enthalten sind, um die geforderte Wahrscheinlichkeit bei dieser Auswahl von Punkten zu erhalten; offenkundig wird diese Wahrscheinlichkeit von der Größe x und nebstbei

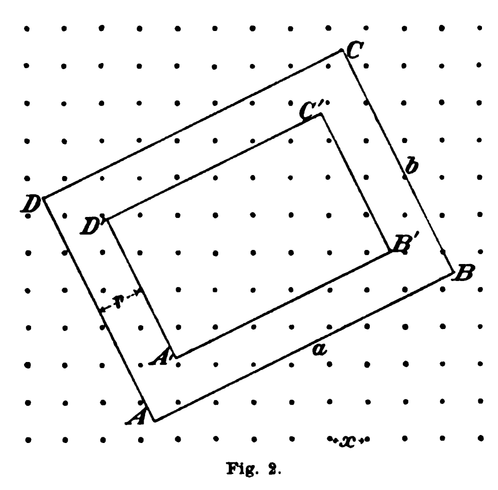


Fig. 2.

---

1) Buffon, Essai d'Arithmétique Morale, 1777. Die Aufgabe wurde aber bereits 1733 der Akademie mitgeteilt.

von der Anordnung der Figur im Punktgitter abhängen. Indem man sich aber vorstellt, daß x immer kleiner und kleiner wird, werden die Anzahlen der Punkte immer mehr den Flächeninhalten der Figuren, auf denen sie liegen, proportional werden und ihr Verhältnis immer weniger beeinflußt sein von der Anordnung der Figur, so daß man das Verhältnis der Gebietsinhalte selbst in dem oben erklärten Sinne als Wahrscheinlichkeit des in Rede stehenden Ereignisses aufzufassen hat; ihr Wert ist sonach $\frac{ab-(a-2r)(b-2r)}{ab}=\frac{2(a+b-2r)r}{ab}$.

Das vorgeführte Beispiel ist geeignet, die Begriffe des *homogenen Gebiets* und der *Dichte* zu vermitteln. Die mit Punkten gleichförmig bedeckte Ebene wie auch das mit Zahlen gleichförmig besetzte Intervall sind homogene Gebiete; ihre Elemente (Punkte, Zahlen) stellen gleich mögliche Fälle dar oder es kommt ihnen, um mit einem vorhin gebrauchten Terminus zu sprechen, gleiche Valenz zu. Es kann aber auch in der Natur des Problems liegen, daß man die Ebene ungleichförmig mit Punkten bedeckt, das Zahlenintervall ungleichförmig mit Zahlen besetzt zu denken hat; dann kommt neben der *Ausdehnung* des Gebiets auch seine variable *Dichte* (Valenz) in Betracht.

Die Außerachtlassung dieses Umstandes hat schon wiederholt zu Fehlschlüssen geführt. Zur schärferen Beleuchtung der Sache sollen zwei Beispiele erörtert werden.[1])

Wenn um die Wahrscheinlichkeit gefragt wird, daß der Ausdruck x^2-yz für willkürliche reelle Werte von x, y, z positiv wird, so ist mit dieser Formulierung gemeint, daß die Bereiche von x, y, z homogen seien.

Bezieht man dieselbe Frage auf $\frac{1}{4}x^2-yz$, so könnte die Meinung entstehen, es müßte dasselbe Resultat herauskommen, da $\frac{1}{2}x$ wie x eine zwischen $-\infty, \infty$ variable Zahl ist: aber diese in der *Ausdehnung* angenommene Identität besteht nicht auch in der *Dichte*. $\frac{1}{4}x^2-yz$ statt x^2-yz nehmen heißt so viel wie stillschweigend voraussetzen, daß x in einem doppelt so dichten Gebiete variieren solle als die beiden andern Variablen.

Daß übrigens die Wahrscheinlichkeit für k^2x^2-yz, wenn x, y, z beliebige positive Zahlen sind, mit k variiert, erkennt man schon daraus, daß sie 0 ist für $k=0$ und sich der 1 nähert, wenn k beständig wächst ($k>0$).

Der Irrtum liegt darin, daß es doch nicht erlaubt ist, statt x willkürlich anzunehmen, kx frei zu wählen und durch k zu dividieren

---

1) Vgl. hierzu E. Cesàro, Considerazioni sul concetto di Probabilità. Periodico di Matematica, VI (1891), p. 1—13, 49—62.

oder x^2 beliebig anzunehmen und daraus die Wurzel zu ziehen. Wenn z. B. bei dem letztgedachten Modus x auf das Intervall $(0, 1)$ angewiesen wird, so ist die Wahrscheinlichkeit, daß es unter $\frac{1}{2}$ bleibt, $\frac{1}{2}$, während sie bei freier Annahme von x^2 nur $\frac{1}{4}$ betrüge. Im ersten Falle ist eben die Dichte unabhängig von x, also gleichförmig, im zweiten ist sie dem x umgekehrt proportional.[1])

Ein Beispiel, aus dem man einen Einwurf gegen dieses Gebiet der Wahrscheinlichkeitsrechnung ableiten wollte, betrifft die nachstehende Schlußfolgerung: „Die Wahrscheinlichkeit, daß das spezifische *Gewicht* einer noch unbekannten Substanz zwischen 7 und 8 oder zwischen 8 und 9 falle, ist beidemal dieselbe; hingegen ist die Wahrscheinlichkeit, daß ihr spezifisches *Volumen* zwischen $\frac{1}{7}$ und $\frac{1}{8}$ oder zwischen $\frac{1}{8}$ und $\frac{1}{9}$ falle, nicht beidemal die gleiche."

Der Schluß stützt sich auf die *gleiche Ausdehnung* der beiden ersten Gebiete, und auf die *ungleiche Ausdehnung* der beiden anderen; wenn man aber der Verteilung der spezifischen Gewichte gleichmäßige Dichte zuschreibt — eine durch nichts gerechtfertigte willkürliche Annahme —, so ist die Verteilung der reziproken Werte, nämlich der spezifischen Volumina, nicht auch gleichmäßig dicht, und daraus entspringt der Trugschluß.

Die korrekte Berücksichtigung dieses Umstandes hebt den scheinbaren Widerspruch auf. Angenommen, die spezifischen Gewichte gingen bis zu einem gewissen Werte, der als Einheit angenommen werden möge. Dann wird die Wahrscheinlichkeit, daß das spezifische Gewicht eines Körpers in das Intervall $(x, x + dx)$ falle, nicht nur von der Ausdehnung dx des Intervalls, sondern auch von seiner durch den Anfangswert gekennzeichneten Lage abhängen und daher durch einen Ausdruck von der Form $\varphi(x)dx$ dargestellt sein, da man annehmen kann, sie sei der Ausdehnung dx proportional. Mit Rücksicht auf die summatorische Bedeutung des Integrals bedeutet (Nr. 30)

$$p = \int_a^b \varphi(x)\,dx$$

die Wahrscheinlichkeit, daß das spezifische Gewicht zwischen a, b falle.

Wenn nun zwei Variable u, v durch eine Gleichung

$$f(u, v) = 0 \tag{4}$$

1) J. Bertrand, Calcul des probabilités, p. 4, vertritt die irrige Anschauung, es sei gleichgültig, ob x direkt oder durch Vermittlung von x^2 willkürlich angenommen wird.

miteinander verbunden sind, und wenn die relative Häufigkeit $\varphi(u)$ der Werte der einen bekannt ist, so folgt daraus für die relative Häufigkeit $\psi(v)$ der Werte der andern zunächst die Beziehung:

$$\varphi(u) : \psi(v) = dv : du;$$

denn je enger das Intervall, auf welches sich die den u in du zugeordneten v zusammendrängen, desto größer ist ihre relative Häufigkeit; aus dem Zusammenhange (4) geht aber anderseits hervor, daß dem Betrage nach

$$dv : du = \frac{\partial f}{\partial u} : \frac{\partial f}{\partial v};$$

folglich gilt:

$$\varphi(u) : \psi(v) = \frac{\partial f}{\partial u} : \frac{\partial f}{\partial v}; \tag{5}$$

woraus sich $\psi(v)$ bestimmen läßt.

Zwischen dem spezifischen Gewichte x und dem spezifischen Volumen y besteht nun die Beziehung

$$xy - 1 = 0;$$

infolgedessen ist

$$\varphi(x) : \psi(y) = y : x,$$

also

$$\psi(y) = \frac{x}{y}\varphi(x) = \frac{1}{y^2}\varphi\left(\frac{1}{y}\right),$$

daher die Wahrscheinlichkeit, das spezifische Volumen liege zwischen $\frac{1}{b}$ und $\frac{1}{a}$,

$$p' = \int_{\frac{1}{b}}^{\frac{1}{a}} \frac{1}{y^2}\varphi\left(\frac{1}{y}\right) dy;$$

durch die Substitution $\frac{1}{y} = x$ geht aber das Integral in $\int_a^b \varphi(x)dx$ über; also ist immer, bei jeder Verteilung der spezifischen Gewichte, $p' = p$, wie es sein muß.[1])

1) Der hier verwendete Gedankengang bildet die Grundlage des Kriteriums, das H. Poincaré, Calcul des probabilités, 1896, p. 97 sq., benützt, um zu beurteilen, ob zwei verschiedene Auffassungen eines Problems gleichberechtigt sind in bezug auf das Resultat. Das n-fache Integral

$$\int_{(n)} \varphi(x_1, x_2, \cdots x_n) dx_1 dx_2 \cdots dx_n$$

49. Die Mengenlehre in der Wahrscheinlichkeitsrechnung. Um einen Überblick über die Probleme der Wahrscheinlichkeitsrechnung zu gewinnen, empfiehlt es sich, sie mit der Mengenlehre in Beziehung zu setzen. Denn die Gesamtheiten der möglichen und der günstigen Fälle lassen sich dem umfassenden und fundamentalen Mengenbegriff unterordnen, und zwar handelt es sich bei den klassischen Aufgaben über Glücksspiele, Ziehungen aus Urnen u. dgl. um *endliche* Mengen, bei den Problemen mit unendlich vielen Fällen um *transfinite* Mengen.

Unter einer *Menge M* versteht man eine Zusammenfassung von irgendwelchen wohlbestimmten und voneinander unterscheidbaren Objekten (Elementen) $A : M = \{A\}$. Für unsere Zwecke kommen in Betracht: Kugeln verschiedener Farbe, Würfelseiten, Kartenblätter, Nummern u. dgl.; Werte einer Variablen, Wertverbindungen mehrerer Variablen, Punkte auf Linien, Flächen, im Raume, Gerade in der Ebene, im Raume, Ebenen u. dgl.

Die Mengen werden unterschieden in endliche und unendliche (*transfinite*).

Dem *Anzahlbegriff* der endlichen Mengen ist übergeordnet der Begriff der *Mächtigkeit*, der bei transfiniten Mengen zur Anwendung kommt. Die Mächtigkeit der Menge M wird durch $\overline{\overline{M}}$ bezeichnet.

Sowie man zwei endliche Mengen als *äquivalent*, d. h. der Anzahl nach gleich, bezeichnet, so nennt man auch transfinite Mengen äquivalent oder von gleicher Mächtigkeit, wenn sich ihre Elemente eindeutig einander zuordnen lassen. Dabei ist als für die transfiniten

---

stellt bei gegebenem φ die Wahrscheinlichkeit dar, daß ein durch die Parameter $x_1, x_2 \cdots x_n$ bestimmtes Gebilde gewissen Bedingungen entspreche; die Integration erstreckt sich dann über jene Wertverbindungen von $x_1, x_2, \cdots x_n$, die aus diesen Bedingungen hervorgehen. Nun werde dasselbe Gebilde mittels anderer Parameter $y_1, y_1, \cdots y_n$ bestimmt, und die Wahrscheinlichkeit für die Einhaltung derselben Bedingungen drücke sich nun durch

$$\int\limits_{(n)} \psi(y_1, y_2, \cdots y_n) dy_1 dy_2 \cdots dy_n$$

aus. Transformiert man dieses zweite Integral mittels der zwischen den beiden Systemen von Parametern bestehenden Relationen in

$$\int\limits_{(n)} \psi \cdot \frac{\partial(y_1, y_2, \cdots y_n)}{\partial(x_1, x_2, \cdots x_n)} dx_1 dx_2, \cdots dx_n$$

und stimmt das Produkt aus ψ und der Jacobischen Determinante $\frac{\partial(y_1, y_2, \cdots y_n)}{\partial(x_1, x_2, \cdots x_n)}$ mit der ursprünglichen Funktion $\varphi(x_1, x_2, \cdots x_n)$ überein, so sind beide Auffassungen äquivalent. — Die von R. Lämmel, l. c., p. 18—19, hierzu gemachte Bemerkung scheint mir unzutreffend.

Mengen charakteristisch hervorzuheben, daß eine solche einer ihrer (ebenfalls transfiniten) Teilmengen äquivalent sein kann. Daß zwei transfinite Mengen M, M' äquivalent seien, wird durch $M \sim M'$ angedeutet.

Als niedrigster Typus einer transfiniten Menge kann die Reihe der natürlichen Zahlen angesehen werden. Sowie die Mächtigkeit einer endlichen Menge durch eine bestimmte Zahl $\mathfrak{e}$ der natürlichen Zahlenreihe ausgedrückt wird, soll die Mächtigkeit dieser selbst durch die ideale Zahl $\mathfrak{a}$ bezeichnet werden (eigentlich unendliche Zahl).

Jede Menge, die der natürlichen Zahlenreihe äquivalent ist, heißt *abzählbar*. Hiernach ist die Reihe der geraden Zahlen abzählbar, weil man ihre Elemente $(2n)$ den Elementen der natürlichen Zahlenreihe (n) eindeutig zuordnen kann; auch die Reihe der echten Brüche ist es, indem man $\frac{1}{2}$ der 2, $\frac{1}{3}, \frac{2}{3}$ der 3, $\frac{1}{4}, \frac{3}{4}$ der 4, $\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$ der 5, usw. zuordnen kann. Daß eine Menge endlich ist, soll durch $\overline{\overline{M}} = \mathfrak{e}$, daß sie transfinit und abzählbar ist, durch $\overline{\overline{M}} = \mathfrak{a}$ bezeichnet werden.

Die Gesamtheit aller reellen Zahlen (und auch die der reellen Zahlen eines Intervalls), ebenso die Gesamtheit der Wertverbindungen zweier und mehrerer stetigen Variablen, geometrisch gesprochen, die Menge der Punkte in einer Geraden, einer Strecke, in der ganzen Ebene oder einem begrenzten Teil derselben, sind *nicht abzählbar;* sie stellen Mengen von höherer Mächtigkeit im Vergleich zur Gesamtheit der natürlichen Zahlen vor. Man spricht hier von der *Mächtigtigkeit eines Kontinuums* $(\mathfrak{C})$. Daß eine Menge M die Mächtigkeit eines Kontinuums besitzt, soll durch $\overline{\overline{M}} = \mathfrak{C}$ angezeigt werden.

Punktmengen P können in bezug auf das *stetige* Gebiet, in dem sie sich befinden (Linie, Fläche, Raum), verschiedene Anordnung zeigen. Ist jeder Punkt der Menge isoliert, d. h. mit einer beliebig engen Umgebung versehen, in der sich kein anderer Punkt der Menge befindet, so heißt die Punktmenge *nirgends dicht.* Gibt es keinen noch so kleinen Gebietsteil, der frei von Punkten der Menge ist, so heißt diese *überall dicht.* Beide Eigenschaften können nebeneinander vorkommen, die Menge ist dann *stellenweise dicht.*

Wenn für die Summe der Gebietsteile, in welchen sich Punkte von P befinden, ein fester Grenzwert existiert, so bezeichnet man ihn als den *Inhalt* von P, symbolisch $J(P)$.

Punktmengen, die Kontinua bilden, sind in ihrem Gebiet überall dicht und man ordnet ihnen den geometrischen Inhalt des Gebietes (Länge der Linie, Größe der Fläche, Volumen des Raumes) als Inhalt zu.

Nach der Natur der Mengen, welche die möglichen und die

günstigen Fälle eines Problems bilden, richtet sich der analytische Vorgang der Wahrscheinlichkeitsbestimmung. Die ursprüngliche Wahrscheinlichkeitsdefinition bezieht sich auf den Fall, daß beide Mengen endlich sind. Bei transfiniten Mengen wird im allgemeinen ein Grenzprozeß zur Wahrscheinlichkeitsbestimmung führen, der von jener Definition den Ausgang nimmt.[1])

50. Geometrische Wahrscheinlichkeiten. Man kann, von der eben erwähnten Systematik Gebrauch machend, als Probleme der geometrischen Wahrscheinlichkeit diejenigen bezeichnen, die unter den Typus (𝔈, 𝔈) fallen, bei denen also mögliche und günstige Fälle Mengen von der Mächtigkeit eines Kontinuums bilden. Aufgaben dieser Art können sowohl arithmetischer wie geometrischer Natur sein. Im ersten Falle handelt es sich um Tatbestände, deren Modalitäten von einer oder mehreren stetig veränderlichen Größen abhängen; die Frage richtet sich nach der Wahrscheinlichkeit, daß ein Tatbestand verwirklicht werde oder verwirklicht sei, für den die Werte der Variablen gewissen Bedingungen genügen. Bei den Problemen geometrischer Natur handelt es sich um die Wahrscheinlichkeit, daß ein aus einer wohldefinierten stetigen Mannigfaltigkeit geometrischer Gebilde nach Belieben herausgegriffenes Individuum gewisse Eigenschaften besitze. Da man Aufgaben der ersten Art dadurch, daß man die auftretenden Variablen entsprechend deutet, häufig auf das geometrische Gebiet übertragen kann, so ist die Bezeichnung „geometrische Wahrscheinlichkeit"[2]) für alle Fragen dieses Typus üblich geworden.

Es ist als ein spezifisches Merkmal der Probleme, denen trans-

---

1) R. Lämmel hat in seiner bereits zitierten Dissertation „Untersuchungen über die Ermittlung von Wahrscheinlichkeiten" den Mengenbegriff systematisch für die Zwecke der Wahrscheinlichkeitsrechnung herangezogen. Indem er die transfiniten Mengen nach den Eigenschaften der Abzählbarkeit und Nichtabzählbarkeit, der Dichte und der Existenz oder Nichtexistenz eines Inhaltes gruppiert, kommt er rein formal (ohne Rücksicht auf die Existenzmöglichkeit) zu 12 Mengentypen, die wieder, indem die möglichen und günstigen Fälle je einer dieser Typen zugezählt werden, zu 144 logisch denkbaren Typen von Wahrscheinlichkeitsproblemen führen. Diese Darstellung leistet vorläufig den Dienst, daß sie geeignet ist, System in die bisher behandelten Fragen zu bringen; nur wenige Typen sind unter diesen Fragen vertreten; manche der Typen lassen sich mangels zureichender Ausbildung der Mengenlehre nicht weiter verfolgen, den meisten dürfte nur zahlentheoretische Bedeutung zukommen; die praktische Tragweite ist sicher nur gering. — Die gewöhnlichen Aufgaben der Wahrscheinlichkeitsrechnung wären unter den oben nicht mitgezählten Typus (𝔢, 𝔢) zu stellen, d. h. mögliche und günstige Fälle bilden endliche Mengen. Der niedrigste, deshalb aber nicht etwa am einfachsten zu behandelnde Typus in transfiniten Mengen wäre (𝔞, 𝔞), wo mögliche und günstige Fälle abzählbare Mengen bilden.

2) Local probability, geometrical probability.

finite Mengen zugrunde liegen, hingestellt worden, daß sie häufig verschiedene, einander also anscheinend widersprechende Lösungen zulassen, und man hat dies auf die verschiedene Deutung zurückgeführt, welche den in der Formulierung solcher Probleme auftretenden Redewendungen „willkürlich angenommene Zahl", „willkürlich gewählter Punkt" u. dgl. gegeben werden kann. Diese scheinbare Paradoxie in den Lösungen ist auch als eine mit dem Wesen der Wahrscheinlichkeitsrechnung untrennbar verbunde Erscheinung bezeichnet worden.

Hierzu sei vor allem das Folgende bemerkt. Die Wahrscheinlichkeitsrechnung als solche hat mit diesen Paradoxien nichts zu schaffen. Sie löst ihre Probleme auf Grund bestimmter Urteilsmaterien und dann immer eindeutig; wo der Urteilsmaterie die Bestimmtheit mangelt, ist auch die Basis für eine Wahrscheinlichkeitsberechnung nicht vorhanden.

Der Begriff der Wahrscheinlichkeit eines Ereignisses oder Tatbestandes *an sich*, ohne Bezugnahme auf eine bestimmte Materie, entbehrt jeder Grundlage. Von der Wahrscheinlichkeit einer weißen Kugel zu sprechen hat keinen Sinn, wofern man sich nicht etwa unter Berufung auf das absolute Nichtwissen auf das Prinzip des mangelnden Grundes stützen und jene Wahrscheinlichkeit mit $\frac{1}{2}$ bewerten will. Ist hingegen die Frage in bezug auf eine bestimmte Kugelmenge gestellt, in der sich weiße Kugeln befinden, und besteht Gleichmöglichkeit der Fälle oder wird sie angenommen, so gibt die Theorie eine bestimmte und nur eine Antwort.

In gleicher Weise verhält es sich mit den Problemen über transfinite Mengen. Ist die Struktur der Menge bestimmt, so ergibt sich eine eindeutige Lösung. Wo über diese Struktur verschiedene Annahmen gemacht werden, ergeben sich auch mehrere im allgemeinen verschiedene Lösungen. Wenn z. B., um auf ein später zu behandelndes Problem vorzubereiten, die Wahrscheinlichkeit verlangt wird, daß eine Sehne des Kreises länger sei als die Seite des ihm eingeschriebenen regulären Dreiecks, so kommt es darauf an, welcher transfiniten Mannigfaltigkeit man die Sehne zuschreibt: ob man die Sehne auffaßt als den Abschnitt, den der Kreis auf einer in der Ebene *beliebig gezogenen Geraden* bildet, oder als die Verbindungslinie zweier auf dem Umfang *beliebig angenommenen Punkte* o. dgl. Es sind ganz verschiedene Mannigfaltigkeiten, zu denen diese Auffassungen führen, aber auch erst dann führen, wenn man darüber einig geworden ist, was man unter einer in der Ebene beliebig gezogenen Geraden, beziehungsweise unter einem auf dem Kreise beliebig angenommenen Punkte zu verstehen habe.

Wie man sieht, führt die Analyse zu der Notwendigkeit gewisser *Grundannahmen* hin, von denen dann bei Behandlung jedes Problems naturgemäßer Gebrauch zu machen ist; diese Grundannahmen sollen in den einen Satz zusammengefaßt werden: Ein Element (Zahl, Punkt, Gerade o. dgl.) schlechtweg willkürlich annehmen heißt, die transfinite Menge, also das Kontinuum, dem es angehört, als *homogen* voraussetzen; dies bedeutet so viel als: die Wahrscheinlichkeit, das Element gehöre einem bestimmten Teil des Kontinuums an, lediglich von der Ausdehnung, dem Inhalt dieses Teils und nicht auch von seiner Lage im Gesamtbereich abhängig sein lassen.

Soll hiernach der Wert der Variablen x in dem Intervall (a, b) willkürlich angenommen werden, so ist die Wahrscheinlichkeit, daß er dem Intervall $(x, x + dx)$ angehöre, durch $\frac{dx}{b-a}$ gegeben. Und sollen die Werte der Variablen x, y je in dem Intervall (a, b), (c, d) willkürlich, also auch unabhängig voneinander gewählt werden, so bestimmt sich die Wahrscheinlichkeit, daß sie gleichzeitig den Intervallen $(x, x + dx)$, $(y, y + dy)$ entnommen seien, durch den Bruch $\frac{dx\,dy}{(b-a)(d-c)}$.

Ist nun ein geometrisches Gebilde als durch die Werte dieser Variablen in den genannten Intervallen bestimmt anzusehen, so ist die transfinite Menge der möglichen Gestaltungen des Gebildes auf das homogene Feld der Wertverbindungen

$$x \mid y; \qquad a \leqq x \leqq b, \qquad c \leqq y \leqq d$$

abgebildet; und das Integral

$$\iint \frac{dx\,dy}{(b-a)(d-c)},$$

ausgedehnt über alle jene Wertverbindungen, denen Gebilde mit der Eigenschaft entsprechen, nach deren Wahrscheinlichkeit gefragt wird, liefert den numerischen Wert dieser Wahrscheinlichkeit.

Man kann den Problemen der Wahrscheinlichkeitsrechnung gegenüber, soweit sie sich auf transfinite Mengen beziehen, zweierlei Standpunkt einnehmen. Man kann fordern, daß die Formulierung so detailliert sei, daß über die Struktur der Mengen kein Zweifel übrig bleibt. Man kann aber auch den Grundsatz befolgen, die dem Wortlaute des Problems am natürlichsten sich anschließende Lösung zu suchen, und im übrigen an der Grundannahme der homogenen Bereiche festhalten. Unter allen Umständen ist, sobald man an die Wirklichkeit herantritt, genau darauf zu achten, daß die Realisierung

der allgemeinen Bedingungen den bei der Lösung des Problems getroffenen Annahmen entspreche.

Nach dem Vorgeführten käme die Berechnung geometrischer Wahrscheinlichkeiten in letzter Linie auf die Auswertung bestimmter Integrale zurück. Aber schon bei verhältnismäßig einfachen Problemen gestaltet sich dieser Weg, insbesondere rücksichtlich der Feststellung der Integrationsgrenzen, oftmals recht schwierig. Durch Ausbildung scharfsinniger Methoden für manche häufig wiederkehrende Bestimmungen von Mannigfaltigkeitsinhalten, durch Zurückführung komplizierter Fragen auf einfache Grundprobleme u. ä. ist es gelungen, die Integrationen mitunter völlig zu umgehen; indirekt sind auf diesem Wege manche schwierige Integralbestimmungen gelungen. Hauptsächlich waren es englische Geometer, die diesen Zweig der Wahrscheinlichkeitsrechnung gefördert haben. Eine Auswahl von Beispielen[1]) wird am besten einen Einblick in seine Methodik vermitteln. Die Beschäftigung mit derlei Aufgaben hat auch einen hohen didaktischen Wert, indem sie für andere Anwendungsgebiete die zweckmäßigste Vorbereitung gewährt.

51. Beispiel XXXI. *Mit welcher Wahrscheinlichkeit läßt sich aus drei Strecken x, y, z, die unter einer gemeinsamen oberen Grenze liegen, ein Dreieck bilden?*

Faßt man x, y, z als (rechtwinklige) Koordinaten eines Punktes im Raume auf, so ist das Gebiet der möglichen Fälle durch den Würfel OD, Fig. 3, von der Kante a dargestellt; denn für Punkte innerhalb oder an der Oberfläche dieses Würfels — und nur für solche — ist gleichzeitig

$$0 \leqq x \leqq a, \qquad 0 \leqq y \leqq a, \qquad 0 \leqq z \leqq a.$$

Soll ein Dreieck gebildet werden können, so muß

$$y + z \geqq x, \qquad z + x \geqq y, \qquad x + y \geqq z$$

Fig. 3.

sein; die Gebiete, welche diesen Bedingungen entsprechen, werden von dem Würfel durch die Ebenen OBC, OCA, OAB abgeschnitten und bilden drei kongruente Pyramiden, $OBCD$, $OCAD$, $OABD$. Das Verhältnis ihres Inhaltes zum Inhalt des Würfels gibt die verlangte Wahrscheinlichkeit:

1) In betreff weiterer Beispiele über „Geometrische Wahrscheinlichkeiten und Mittelwerte" kann auf des Verfassers gleichnamige Schrift (Leipzig 1884; in französischer Übersetzung Paris 1902 erschienen) verwiesen werden.

$$p = \frac{3\,\frac{a^3}{6}}{a^3} = \frac{1}{2}.$$

52. Beispiel XXXII. *Eine Strecke a wird beliebig in drei Teile geteilt; wie groß ist die Wahrscheinlichkeit, daß sich aus den Teilen ein Dreieck bilden lasse?*

Die Teile x, y, z müssen positiv sein und der Gleichung

$$x + y + z = a \tag{1}$$

genügen.

Damit sich aus ihnen ein Dreieck bilden lasse, müssen sie den Bedingungen

$$y + z \geqq x, \qquad z + x \geqq y, \qquad x + y \geqq z$$

genügen, welche sich mit Hilfe der vorangehenden Gleichung umwandeln lassen in

$$x \leqq \frac{a}{2}, \qquad y \leqq \frac{a}{2}, \qquad z \leqq \frac{a}{2}. \tag{2}$$

Werden x, y, z als rechtwinklige Koordinaten eines Punktes im Raume aufgefaßt, so ist das Gebiet der möglichen Fälle durch das Dreieck ABC, Fig. 4, dargestellt, dessen Ecken von O den Abstand a haben; denn die Koordinaten jedes Punktes in diesem Dreieck sind positiv und genügen der Gleichung (1). Das Gebiet der günstigen Fälle wird aus diesem Dreieck durch die drei Ebenen EFG, HFJ, KJG ausgeschnitten, welche OA, beziehungsweise OB, OC senkrecht halbieren; es ist dies das Dreieck JGF, dessen Punkte den Beziehungen (2) entsprechen.

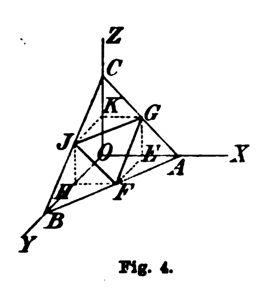


Fig. 4.

Da nun $\triangle JGF = \frac{1}{4}\triangle ABC$ ist, so ist die verlangte Wahrscheinlichkeit

$$p = \frac{1}{4}.$$

Man kann, von der Tatsache Gebrauch machend, daß die Summe der Lote, welche von einem Punkte in der Fläche eines gleichseitigen Dreiecks auf dessen Seiten gefällt werden, konstant und gleich der Höhe des Dreiecks ist, die Aufgabe auch in der Ebene erledigen.

53. Beispiel XXXIII. *Auf eine horizontale, mit äquidistanten Parallelen vom Abstande 2a überzogene Tafel wird eine (zylindrische)*

Nadel von der Länge $2c$ ($\leq 2a$) geworfen; wie groß ist die Wahrscheinlichkeit, daß die Nadel eine der Parallellinien kreuzt?

Es genügt, nur diejenigen Lagen der Nadel ins Auge zu fassen, bei welchen ihr Mittelpunkt M, Fig. 5, mit einem Punkt von $AC = \frac{AB}{2} = a$ zusammenfällt; die Menge solcher Lagen wird durch πa gemessen; denn a ist das Maß für die Menge der Punkte in AC und π das Maß für die Menge der Richtungen bei jeder Lage des Mittelpunktes.

Fällt der Mittelpunkt in die Entfernung $x = AM$ bis $x + dx$, so ist die Menge der günstigen Nadelrichtungen bei jeder solchen Lage des Mittelpunktes durch 2θ, d. i.

$$2 \arccos \frac{x}{c}$$

gemessen; die Menge der beschriebenen Lagen des Mittelpunktes hat aber dx zum Maße; somit ist das Gebiet der günstigen Fälle durch

$$\int_0^c 2 \arccos \frac{x}{c}\, dx = 2c$$

gemessen.

Die verlangte Wahrscheinlichkeit ist demnach

$$p = \frac{2c}{\pi a}.$$

Die vorstehende Aufgabe, unter dem Namen „Nadelproblem" (problème de l'aiguille) bekannt, gehört zu den ersten, welche auf dem Gebiete der geometrischen Wahrscheinlichkeit gestellt worden sind; ihr Urheber ist Buffon[1]). Über ihre weitere Entwicklung vergleiche man des Verfassers „Geometrische Wahrscheinlichkeiten und Mittelwerte" (1884), p. 85 ff. und dessen Bericht über „Die Entwicklung der Wahrscheinlichkeitstheorie" (1899), VII. Jahresber. der Deutschen Mathematiker-Verein., p. 59 ff.

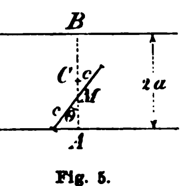


Fig. 5.

34. Beispiel XXXIV. *In einem Kreise werden, nachdem darin eine Sehne beliebig gezogen worden, zwei Punkte willkürlich angenommen. Mit welcher Wahrscheinlichkeit liegen beide Punkte zu einerlei Seite der Sehne?*

α) Wird die Sehne bei beliebig angenommener Richtung in beliebiger Entfernung vom Kreismittelpunkt gezogen, so ist die Menge

1) B.s sämtliche Werke, deutsch von B. Rave, 4. Bd., p. 441—498.

der Sehnen, deren Entfernung zwischen $x = OC$, Fig. 6, und $x + dx$ liegt, durch dx bestimmt. Die Menge der Punktepaare, welche bei dieser Lage der Sehne AB unterhalb derselben liegen, wird durch das Quadrat des Segments ABD, die Menge der Punktepaare oberhalb AB durch das Quadrat des Segments ABE gemessen. Hiernach ergibt sich das Maß für die Mannigfaltigkeit der günstigen Fälle, indem man den Ausdruck

$$\left\{(ABD)^2 + (ABE)^2\right\} dx$$

Fig. 6.

über alle Sehnenlagen integriert; nach Einführung des Winkels $AOC = \varphi$ und des Kreishalbmessers $AO = r$ ergibt dies:

$$\int_0^\pi \left\{r^4 (\varphi - \sin\varphi \cos\varphi)^2 + r^4 (\pi - \varphi - \sin\overline{\pi - \varphi} \cos\overline{\pi - \varphi})^2\right\} r \sin\varphi d\varphi$$

und mit Rücksicht darauf, daß

$$\int_0^\pi (\pi - \varphi - \sin\overline{\pi - \varphi} \cos\overline{\pi - \varphi})^2 \sin\varphi d\varphi = \int_0^\pi (\varphi - \sin\varphi \cos\varphi)^2 \sin\varphi d\varphi,$$

weiter

$$2r^5 \int_0^\pi (\varphi - \sin\varphi \cos\varphi)^2 \sin\varphi d\varphi = 2r^5 \left(\pi^2 - \frac{128}{45}\right).$$

Die Menge der möglichen Fälle hat $2r \cdot (\pi r^2)^2$ zum Maße. Hiernach ist

$$p = 1 - \frac{128}{45\pi^2} = 0{,}712 \cdots$$

die verlangte Wahrscheinlichkeit.

β) Zu einem andern Resultate gelangt man, wenn man sich die Sehne aus dem freigewählten Anfangspunkt A in beliebiger Richtung gezogen denkt. Denn die Menge solcher Sehnen, deren Lot OC mit AO einen zwischen φ und $\varphi + d\varphi$ liegenden Winkel einschließt, ist durch $d\varphi$, daher die Menge der günstigen Fälle durch

$$\int_0^\pi \left\{r^4 (\varphi - \sin\varphi \cos\varphi)^2 + r^4 (\pi - \varphi - \sin\overline{\pi - \varphi} \cos\overline{\pi - \varphi})^2\right\} d\varphi$$

$$= 2r^4 \left(\frac{\pi^3}{3} + \frac{3\pi}{8} + \frac{1}{8}\right)$$

auszudrücken, während die Menge der möglichen Fälle nunmehr $\pi \cdot (\pi r^3)^2$ zum Maße hat.

Bei dieser Auffassung ist also

$$p' = \frac{2}{3} + \frac{3\pi + 1}{4\pi^3} = 0{,}747\cdots$$

die gesuchte Wahrscheinlichkeit.

55. Mittel- oder Durchschnittswert einer vom Zufall abhängigen Größe. Eine veränderliche Größe S sei der Werte S_1, $S_2, \cdots S_\nu$ fähig, deren jeden sie nur auf eine Art annehmen kann. Das arithmetische Mittel dieser Einzelwerte wird als *Mittel- oder Durchschnittswert von* S bezeichnet. Bedient man sich dafür des Symbols $\mathfrak{D}(S)$, so ist

$$\mathfrak{D}(S) = \frac{S_1 + S_2 + \cdots + S_\nu}{\nu}. \tag{1}$$

Kann jeder Einzelwert von S auf mehrere Arten zustande kommen, allgemein S_i auf n_i Arten, und setzt man $n_1 + n_2 + \cdots + n_\nu = N$, so ist

$$\mathfrak{D}(S) = \frac{n_1 S_1 + n_2 S_2 + \cdots + n_\nu S_\nu}{N}. \tag{2}$$

Unterliegt die Größe S der willkürlichen Annahme und kann jeder ihrer Einzelwerte mit gleichem Grade der Möglichkeit aus der Wahl hervorgehen, so ist im ersten Falle $\frac{1}{\nu}$, im zweiten Falle $\frac{n_1}{N} = p_1$ die Wahrscheinlichkeit, daß sich der Wert S_1 einstellen werde, ebenso $\frac{1}{\nu}$, beziehungsweise $\frac{n_2}{N} = p_2$ die Wahrscheinlichkeit des Wertes S_2 usw. Mithin ist unter diesem Gesichtspunkte

$$\mathfrak{D}(S) = p_1 S_1 + p_2 S_2 + \cdots + p_\nu S_\nu, \tag{3}$$

d. h. *der Mittelwert der vom Zufall abhängigen Größe S ist gleich der Summe ihrer Einzelwerte, jeder mit der seinem Zustandekommen entsprechenden Wahrscheinlichkeit multipliziert.*

Ist die Mannigfaltigkeit der Werte S eine stetige, indem der einzelne von der Wertverbindung der stetigen Variabeln $x, y, \cdots$ abhängt, so ist die Menge der Werte S, welche zu Wertverbindungen zwischen $x, y, \cdots$ und $x + dx$, $y + dy, \cdots$ gehören, gemessen durch $dx dy \cdots$, die Wahrscheinlichkeit eines solchen Wertes von S gleich

$$\frac{dx\,dy\cdots}{\int\int\cdots dx\,dy\cdots},$$

das Integral über die ganze Mannigfaltigkeit ausgedehnt. Der Mittelwert von S ist dann in Konsequenz des voranstehenden Satzes

$$\mathfrak{D}(S) = \frac{\iint \cdots S\,dx\,dx \cdots}{\iint \cdots dx\,dy \cdots}, \tag{4}$$

beide Integrationen über das ganze Gebiet der S erstreckt.

Es sei bemerkt, daß man $\mathfrak{D}(S)$ auch häufig als den wahrscheinlichen Wert von S bezeichnet.

56. Zusammenhang zwischen Wahrscheinlichkeit und Mittelwert. Der grundlegende, diesen Zusammenhang betreffende Satz ist der folgende: „Ist das variable geometrische Gebiet S (Linie, Fläche, Raum) in dem festen gleichartigen Gebiet A eingeschlossen, $\mathfrak{D}(S)$ der Mittelwert von S, so ist die Wahrscheinlichkeit, daß ein in A willkürlich angenommener Punkt in S falle, welchen Wert dieses im Augenblicke der Realisierung auch haben möge,

$$p = \frac{\mathfrak{D}(S)}{A}.\text{“}$$

Kann nämlich S die Werte $S_1, S_2, \cdots S_\nu$ mit den Wahrscheinlichkeiten $w_1, w_2, \cdots w_\nu$ annehmen, so ist $w_i \frac{S_i}{A}$ die zusammengesetzte Wahrscheinlichkeit dafür, daß bei der Realisierung der Bedingungen der Wert S_i existiere und daß der im Gebiete A angenommene Punkt in das Teilgebiet S falle. Die totale Wahrscheinlichkeit des letzten Ereignisses ist daher

$$p = w_1 \frac{S_1}{A} + w_2 \frac{S_2}{A} + \cdots + w_\nu \frac{S_\nu}{A} = \frac{w_1 S_1 + w_2 S_2 + \cdots + w_\nu S_\nu}{A} = \frac{\mathfrak{D}(S)}{A}.$$

Die Ausdehnung des Beweises auf eine stetige Mannigfaltigkeit der S unterliegt keiner Schwierigkeit.

Würde es sich statt eines Punktes um deren n handeln, so ergäbe sich für die Wahrscheinlichkeit, daß sie alle in das eingeschlossene Gebiet S fallen, der Ausdruck

$$p = \frac{\mathfrak{D}(S^n)}{A^n}.$$

Die vorstehenden Formeln können zweifache Verwendung finden: zur Berechnung von p oder zur Berechnung von $\mathfrak{D}(S)$, respektive $\mathfrak{D}(S^n)$.

57. Beispiel XXXV. *In einer Strecke von der Länge a werden zwei Punkte X, Y willkürlich angenommen; welches ist der Mittelwert ihres Abstandes $\overline{XY} = s$, welches der Mittelwert seiner n-ten Potenz?*

Denkt man sich auf der Strecke noch einen dritten Punkt Z angenommen, so ist die Wahrscheinlichkeit, daß er zwischen X und Y falle:

$$p = \frac{\mathfrak{D}(s)}{a}.$$

Nun kann p auf kombinatorischem Wege gefunden werden. Die drei Punkte können mit gleichem Grade der Möglichkeit eine der folgenden Anordnungen zeigen:

$$XYZ, \quad XZY, \quad YXZ, \quad YZX, \quad ZXY, \quad ZYX;$$

nur zwei davon, die zweite und vierte, sind dem bezeichneten Ereignis günstig, seine Wahrscheinlichkeit ist daher $p = \frac{1}{3}$. Daraus folgt

$$\mathfrak{D}(s) = \frac{1}{3} a.$$

Um den Mittelwert von s^n zu finden, denke man sich nach X, Y noch n weitere Punkte angenommen; die Wahrscheinlichkeit, daß sie sämtlich zwischen X und Y fallen, ist

$$p = \frac{\mathfrak{D}(s^n)}{a^n}.$$

Sie kann aber auch direkt bestimmt werden. Jeder von den $n + 2$ Punkten kann bei einer bestimmten Zählungsrichtung der erste sein; daß es X *oder* Y sei, hat die Wahrscheinlichkeit $\frac{2}{n+2}$; ist eines von beiden eingetroffen, so ist die Wahrscheinlichkeit, daß der andere Punkt (Y oder X) unter den übrigen der letzte sei, $\frac{1}{n+1}$. Hiernach ist auch

$$p = \frac{2}{(n+2)(n+1)},$$

folglich

$$\mathfrak{D}(s^n) = \frac{2}{(n+2)(n+1)} a^n.$$

58. Erster Satz über Mittelwerte. „Es seien A und B zwei einander ausschließende Gebiete gleicher Dimensionszahl; S sei eine Größe, deren Wert von zwei willkürlich angenommenen Punkten X, Y abhängt; ferner bedeute

μ den Mittelwert von S, wenn ein Punkt in A, der andere in B angenommen wird;

μ_A den Mittelwert, wenn beide Punkte in A,
μ_B den Mittelwert, wenn beide Punkte in B,
μ' den Mittelwert, wenn beide Punkte ohne Wahl im Gesamtgebiete $A + B$ angenommen werden. Dann besteht die Beziehung:

$$(A + B)^2 \mu' = A^2 \mu_A + B^2 \mu_B + 2AB\mu.\text{“}$$

Man denke sich in der Gesamtheit der Punktepaare in $A + B$ zu jedem das zugehörige S; dann zerfällt die Summe Σ aller S in folgende Bestandteile: in Σ_A, das sich auf Punktepaare in A; in Σ_B, das sich auf Punktepaare in B bezieht; endlich in die einander gleichen Summen Σ_{AB} und Σ_{BA}, welche den Fällen entsprechen, wo X in A, Y in B und umgekehrt liegt. Hiernach ist also

$$\Sigma = \Sigma_A + \Sigma_B + 2\Sigma_{AB};$$

nun ist aber

$$\frac{\Sigma}{(A+B)^2} = \mu', \quad \frac{\Sigma_A}{A^2} = \mu_A, \quad \frac{\Sigma_B}{B^2} = \mu_B, \quad \frac{\Sigma_{AB}}{AB} = \mu,$$

folglich in der Tat

$$(A + B)^2 \mu' = A^2 \mu_A + B^2 \mu_B + 2AB\mu.$$

59. Zweiter Satz über Mittelwerte. „Sind A, B, C drei einander ausschließende Gebiete in einer Ebene (Linie oder Fläche) von solcher Beschaffenheit, daß keine Gerade alle drei zugleich schneidet; nimmt man in denselben die Punkte X, Y, Z willkürlich an, so ist die mittlere Fläche des Dreiecks XYZ gleich der Fläche des Dreiecks der Schwerpunkte von A, B, C.“

Der Beweis dieses Satzes beruht auf der Tatsache, daß die mittlere Entfernung des Punktes eines ebenen Gebiets von einer in derselben Ebene gelegenen, das Gebiet nicht schneidenden Geraden gleich ist der Entfernung seines Schwerpunktes von dieser Geraden.

Daraus ergibt sich, daß der Mittelwert aller Dreiecke XYZ mit festem Y, Z durch das Dreieck $\mathfrak{A}YZ$ dargestellt wird, wenn $\mathfrak{A}$ den Schwerpunkt von A bedeutet; weiter ist der Mittelwert aller Dreiecke $\mathfrak{A}YZ$ mit festem Z durch das Dreieck $\mathfrak{A}\mathfrak{B}Z$ bestimmt, wenn $\mathfrak{B}$ der Schwerpunkt von B ist; endlich ist $\mathfrak{A}\mathfrak{B}\mathfrak{C}$ der Mittelwert der Dreiecke $\mathfrak{A}\mathfrak{B}Z$, also zugleich aller Dreiecke XYZ, wenn $\mathfrak{C}$ den Schwerpunkt von C bedeutet.

60. Beispiel XXXVI. *In der Fläche eines gegebenen Dreiecks ABC werden zwei Punkte X, Y beliebig angenommen; welches ist die mittlere Fläche des Dreiecks XYC?*

Teilt man das Dreieck durch die Mittellinie CD, Fig. 7, in zwei gleiche Dreiecke, so ist die mittlere Fläche von XYC *für den Fall, daß* in jede Hälfte einer der Punkte X, Y fällt, gleich dem Dreieck welches die Schwerpunkte von ADC, DBC mit C verbindet; dieses Dreieck hat aber, weil seine Basis $\frac{1}{3}$ von AB und seine Höhe $\frac{2}{3}$ der Höhe von ABC ist, die Fläche $\frac{2}{9} ABC = \frac{2}{9} \Delta$.

Fig. 7.

Nun kann der Satz in Nr. 57 zur Anwendung gebracht werden, indem man

$$A + B = \Delta, \qquad A = B = \frac{1}{2} \Delta, \qquad \mu = \frac{2}{9} \Delta$$

setzt und beachtet, daß $\mu_A = \mu_B = \frac{1}{2} \mu'$ ist, weil ein Dreieck XYC in ADC oder in DBC die halbe mittlere Fläche hat von derjenigen, welche einem solchen Dreieck in ABC zukommt; mithin ist

$$\Delta^2 \mu' = \frac{1}{8} \Delta^2 \mu' + \frac{1}{8} \Delta^2 \mu' + \frac{1}{9} \Delta^3,$$

woraus der gesuchte Mittelwert

$$\mu' = \frac{4}{27} \Delta$$

folgt.

61. Beispiel XXXVII. *In einem Dreieck ABC werden drei Punkte X, Y, Z beliebig angenommen; es ist die Wahrscheinlichkeit zu bestimmen, daß sie mit der Ecke C ein konvexes Viereck ergeben.*

Die Punkte ergeben ein nichtkonvexes Viereck, wenn z. B. der Punkt Z in das Dreieck XYC zu liegen kommt; die Wahrscheinlichkeit hierfür ist, wenn μ' den Mittelwert von XYC und Δ die Fläche von ABC bezeichnet, $\frac{\mu'}{\Delta}$, d. i. $\frac{4}{27}$.

Da nun auch Y in das Dreieck ZXC oder X in das Dreieck YZC fallen kann, so ist die totale Wahrscheinlichkeit eines nichtkonvexen Vierecks $\frac{4}{9}$, folglich die Wahrscheinlichkeit eines konvexen $\frac{5}{9}$.

62. Beispiel XXXVIII. *In einem Halbkreise vom Halbmesser a werden zwei Punkte X, Y beliebig angenommen und mit einem Endpunkte des begrenzenden Durchmessers zu einem Dreieck verbunden; es ist die mittlere Fläche dieses Dreiecks zu bestimmen.*

Bestimmt man die Punkte X, Y durch ihre Polarkoordinaten r, φ, beziehungsweise r', φ' in bezug auf den dritten Eckpunkt A

des Dreiecks, Fig. 8, als Pol und den Durchmesser AB als Polarachse, so ist das Element der Ebene bei X durch $r\,dr\,d\varphi$, bei Y durch $r'\,dr'\,d\varphi'$ bestimmt, und die Menge der Dreiecke, deren Ecken X, Y in diese Elemente fallen, durch $r\,dr\,d\varphi\,r'\,dr'\,d\varphi'$ gemessen; da jedes solche Dreieck mit der Fläche $\frac{1}{2}\,rr'\sin(\varphi'-\varphi)$ in Rechnung zu stellen ist, so hat man

$$\iiiint r^2 r'^2 \sin(\varphi'-\varphi)\,dr\,dr'\,d\varphi\,d\varphi' \tag{1}$$

als Ausdruck für die Summe aller Dreiecksflächen, wenn die Integration so geführt wird, daß $\varphi' > \varphi$ ist. Nach Vollziehung der auf r, r' bezüglichen Integrationen, deren Grenzen $0, 2a\cos\varphi$, respektive $0, 2a\cos\varphi'$ sind, geht dies über in

Fig. 8.

$$\frac{64a^6}{9}\int_0^{\frac{\pi}{2}} d\varphi' \int_0^{\varphi'} \cos^3\varphi \cos^3\varphi' \sin(\varphi'-\varphi)\,d\varphi$$

$$= \frac{64a^6}{9}\left\{\int_0^{\frac{\pi}{2}} \cos^3\varphi' \sin\varphi'\,d\varphi' \int_0^{\varphi'} \cos^4\varphi\,d\varphi - \int_0^{\frac{\pi}{2}} \cos^4\varphi'\,d\varphi' \int_0^{\varphi'} \cos^3\varphi \sin\varphi\,d\varphi\right\};$$

wendet man auf das erste Integral die Transformation[1])

$$\int_0^a dx \int_0^x f(x, y)\,dy = \int_0^a dx \int_0^x f(a-y, a-x)\,dy$$

an, so verwandelt es sich in das Integral

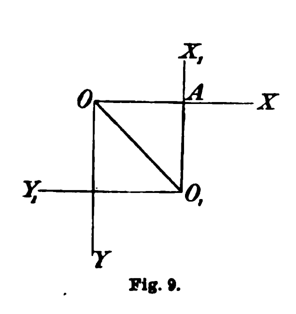


Fig. 9.

1) Das Integral $\int_0^a dx \int_0^x f(x, y)\,dy$ erstreckt sich über das gleichschenklige Dreieck OAO_1, Fig. 9; geht man von dem Koordinatensystem XOY zu dem System $X_1\,O_1\,Y_1$ über, so lauten die Transformationsgleichungen:

$$x = a - y_1, \qquad y = a - x_1,$$

und das transformierte Integral ist

$$\int_0^a dx_1 \int_0^x f(a - y_1, a - x_1)\,dy_1.$$

$$\int_0^{\frac{\pi}{2}} \sin^4\varphi'\,d\varphi' \int_0^{\varphi'} \sin^3\varphi \cos\varphi\,d\varphi = \frac{1}{4}\int_0^{\frac{\pi}{2}} \sin^8\varphi'\,d\varphi';$$

das zweite Integral hingegen ist gleich

$$\frac{1}{4}\int_0^{\frac{\pi}{2}} \cos^4\varphi'\,d\varphi' \left\{\cos^4\varphi\right\}_0^{\varphi'} = \frac{1}{4}\int_0^{\frac{\pi}{2}} \cos^8\varphi'\,d\varphi' - \frac{1}{4}\int_0^{\frac{\pi}{2}} \cos^4\varphi'\,d\varphi';$$

demnach hat (1) den Wert

$$\frac{64a^6}{9\cdot 4}\left\{\int_0^{\frac{\pi}{2}} \sin^8\varphi'\,d\varphi' + \int_0^{\frac{\pi}{2}} \cos^8\varphi'\,d\varphi' - \int_0^{\frac{\pi}{2}} \cos^4\varphi'\,d\varphi'\right\} = \frac{11a^6}{4\cdot 18}.$$

Da die Menge der Dreiecke, übereinstimmend mit der Menge der Punktepaare in der Halbkreisfläche, durch $\left(\frac{\pi a^2}{2}\right)^2$ gemessen wird, so ist die verlangte mittlere Fläche

$$\mathfrak{D}(AXY) = \frac{11a^2}{18\pi}.$$

63. Dritter Satz über Mittelwerte. „Ist μ der Mittelwert einer Größe S, welche von n in dem ebenen Gebiete A beliebig angenommenen Punkten abhängt, μ_1 der Mittelwert derselben Größe für den Fall, daß einer der Punkte an der Begrenzung von A angenommen wird; ändert man ferner A um den differentiellen Betrag dA und bezeichnet die Änderung von μ mit $d\mu$, so besteht zwischen diesen Größen die Differentialgleichung

$$A\,d\mu = n(\mu_1 - \mu)\,dA.“$$

Alle Größen S, welche zu Punktsystemen in dem erweiterten Gebiete $A + dA$ gehören, ergeben die Summe

$$(\mu + d\mu)(A + dA)^n, \tag{1}$$

d. i. das Produkt aus ihrem Mittelwerte mit der Menge der Punktsysteme.

Die Summe zerfällt in die Summe solcher S, welche zu Punktsystemen in A gehören, — diese Summe beträgt

$$\mu A^n; \tag{2}$$

— und in die Summe solcher S, welche zu Punktsystemen gehören, die aus $n - 1$ Punkten in A und einem Punkt in dA bestehen, — diese Summe beträgt

$$n\mu_1 A^{n-1} dA. \tag{3}$$

Bleibt man bei Größen von der Ordnung dA stehen, so ist damit die Summe aller S erschöpft; denn Punktsysteme, in welchen mehr als ein Punkt auf dA entfällt, führen zu Gliedern höherer Ordnung in bezug auf dA.

Entwickelt man auch den Ausdruck (1) bis auf Glieder der ersten Ordnung und setzt ihn dann der Summe der Ausdrücke (2) und (3) gleich, so ergibt sich:

$$\mu A^n + A^n d\mu + n\mu A^{n-1} dA = \mu A^n + n\mu_1 A^{n-1} dA,$$

woraus nach entsprechender Reduktion

$$A d\mu = n(\mu_1 - \mu) dA$$

folgt.

Mittels dieser Gleichung kann aus bekanntem μ_1, dessen Bestimmung sich mitunter einfacher gestaltet als die direkte Berechnung von μ, dieses letztere ermittelt werden. Durch eine zweckmäßige Wahl von dA, welche die Angabe des Verhältniswertes $\frac{d\mu}{dA}$ zuläßt, kann unter Umständen die Integration erspart werden.

Zur Erläuterung diene die folgende, in Nr. 57 bereits gelöste Aufgabe: Auf einer geraden Strecke von der Länge a werden zwei Punkte X, Y willkürlich angenommen; es ist der Mittelwert μ der n-ten Potenz von $XY = s$ zu bestimmen. — Läßt man X mit einem Endpunkt der Strecke zusammenfallen, so wird μ_1 der Mittelwert der n-ten Potenz des Abstandes von Y von jenem Endpunkt; die Wahrscheinlichkeit, daß n weitere Punkte innerhalb dieses Abstandes zu liegen kommen, ist einerseits $\frac{\mu_1}{a^n}$, andererseits gleichbedeutend mit der Wahrscheinlichkeit, daß Y unter $n+1$ Punkten der letzte sei; folglich ist $\frac{\mu_1}{a^n} = \frac{1}{n+1}$, woraus $\mu_1 = \frac{a^n}{n+1}$. Läßt man nun a um da wachsen, so ergibt sich zur Bestimmung von μ die Differentialgleichung

$$a d\mu = 2\left(\frac{a^n}{n+1} - \mu\right) da,$$

welche sich auf die Normalform der linearen Gleichung:

$$\frac{d\mu}{da} + \frac{2}{a}\mu = \frac{2a^{n-1}}{n+1}$$

bringen läßt; ihr allgemeines Integral ist

$$a^2\mu = C + \frac{2a^{n+2}}{(n+1)(n+2)},$$

und da μ mit a zugleich verschwindet, so ist endgültig

$$\mu = \frac{2a^n}{(n+1)(n+2)}.$$

64. Das Vierpunktproblem. *In einer gegebenen ebenen Figur werden vier Punkte beliebig angenommen; wie groß ist die Wahrscheinlichkeit, daß sie sich zu einem konvexen Viereck verbinden lassen?*

Die Aufgabe kommt auf die folgende zurück: Nachdem drei Punkte angenommen worden, die Wahrscheinlichkeit zu bestimmen, daß der vierte Punkt in das durch sie bestimmte Dreieck falle, wodurch ein *nicht*konvexes Viereck zustande kommt. Diese Wahrscheinlichkeit ist aber gleich der mittleren Fläche μ des durch die drei Punkte bestimmten Dreiecks, dividiert durch die Fläche F der Figur. Und da jeder der vier Punkte als der letzte gelten kann, so ist die totale Wahrscheinlichkeit, daß einer von den vier Punkten in das Dreieck aus den drei andern falle, gleich $\frac{4\mu}{F}$, folglich die Wahrscheinlichkeit eines konvexen Vierecks:

$$p = 1 - \frac{4\mu}{F}.$$

Die Schwierigkeit der Lösung des Problems in besonderen Fällen liegt also in der Bestimmung von μ.

Die vorstehende Aufgabe wird[1]) als die erste bezeichnet, welche seit Buffons Nadelproblem auf dem Gebiete der geometrischen Wahrscheinlichkeit gestellt worden ist; ihr Urheber ist J. J. Sylvester.

Spezielle Fälle des Vierpunktproblems, betreffend das Dreieck, das Parallelogramm, das reguläre Sechseck und den Kreis, sind von Woolhouse[2]) und von Crofton[3]) behandelt worden. Daß die Wahrscheinlichkeit p für den Kreis unter allen konvexen Figuren am größten ist, hat der letztere bewiesen; für das Dreieck dürfte sie am kleinsten sein.

In den beiden folgenden Nummern legen wir eine Lösung für das Dreieck und den Kreis vor.

65. Beispiel XXXIX. *Das Vierpunktproblem für das Dreieck zu lösen.*

Es sei $\mathfrak{ABC}$, Fig. 10, das gegebene Dreieck; seine Seiten mögen mit a, b, c bezeichnet werden.

---

1) M. W. Crofton, Probability. Encycl. Brit., 9. edit. (1885), XIX, Art. 69 ff.
2) Educational Times 1867. 3) l. c.

Um den Mittelwert μ eines beliebigen Dreiecks XYZ innerhalb $\mathfrak{ABC}$ zu finden, gehen wir von der spezielleren Aufgabe aus, den Mittelwert μ_1 eines Dreiecks XYZ zu suchen, dessen eine Ecke, etwa Z, auf dem Umfange von $\mathfrak{ABC}$, z. B. in $\mathfrak{AB} = c$ liegt. Zerlegt man $\mathfrak{ABC}$ durch $\mathfrak{C}Z$ in die beiden Teile

$$\mathfrak{AC}Z = A, \qquad \mathfrak{BC}Z = B,$$

Fig. 10.

so daß $\mathfrak{ABC} = A + B$ ist, so ergibt sich μ_1 *bei dieser Lage von* Z mit Hilfe des Satzes in Nr. 58, wenn man kennt:

den Mittelwert eines Dreiecks XYZ in $\mathfrak{AC}Z$,
den Mittelwert eines Dreiecks XYZ in $\mathfrak{BC}Z$,
den Mittelwert eines Dreiecks XYZ in $\mathfrak{ABC}$, wobei X in $\mathfrak{AB}Z$, Y in $\mathfrak{BC}Z$ liegt; nach Nr. 60 ist der erste $\frac{4}{27}A$, der zweite $\frac{4}{27}B$, der dritte nach dem Satze Nr. 59 gleich $\frac{1}{9}(A + B)$. Mithin hat man zur Bestimmung von $\mu_1(Z)$ die Gleichung:

$$(A + B)^2 \mu_1(Z) = \frac{4}{27}A^3 + \frac{4}{27}B^3 + \frac{2}{9}AB(A+B).$$

Daraus berechnet sich zunächst

$$\mu_1(Z) = \frac{2}{27}\left\{2(A + B) - \frac{3AB}{A+B}\right\} = \frac{2}{27}(A + B)\left\{2 - \frac{3}{c^2}x(c - x)\right\},$$

wenn $\mathfrak{A}Z = x$ gesetzt wird; demnach ist der Mittelwert von ZXY, wenn alle Lagen von Z auf $\mathfrak{AB}$ berücksichtigt werden,

$$\mu_1 = \frac{1}{c}\int_0^c \mu_1(Z)\,dx = \frac{1}{9}(A + B) = \frac{1}{9}F,$$

also bloß abhängig von der Fläche F des Dreiecks und daher gültig, wo auch Z auf dessen Umfange angenommen wird.

Zur Bestimmung von μ hat man nun auf Grund des Satzes in Nr. 63 die Differenzialgleichung:

$$F d\mu = 3 dF\left(\frac{1}{9}F - \mu\right);$$

denkt man sich die Erweiterung von F um dF so ausgeführt, daß das Dreieck dabei sich ähnlich bleibt, so ist

$$\frac{\mu + d\mu}{F + dF} = \frac{\mu}{F} = \frac{d\mu}{dF};$$

dadurch verwandelt sich obige Gleichung in

$$\mu = 3\left(\frac{1}{9}F - \mu\right),$$

woraus

$$\mu = \frac{1}{12}F.$$

Nach der allgemeinen Formel der Nr. 64 ist hiernach die Wahrscheinlichkeit eines nichtkonvexen Vierecks $\frac{1}{3}$, die eines konvexen $\frac{2}{3} = 0{,}666\cdots$.

66. Beispiel XI. *Das Vierpunktproblem für den Kreis zu lösen.*

Wir gehen wieder zunächst darauf aus, die mittlere Fläche μ_1 eines Dreiecks XYZ zu bestimmen, von welchem ein Eckpunkt, Z, auf dem Umfange liegt, Fig. 11.

Zerlegt man den Kreis aus Z in die beiden Halbkreise $A = B = \frac{\pi a^2}{2}$, so ist der Mittelwert eines Dreiecks ZXY in A oder in B nach Nr. 62 gleich $\frac{11 a^2}{18\pi}$; der Mittelwert eines Dreiecks ZXY, das, wie das eingezeichnete, einen Eckpunkt in A, den andern in B hat, beträgt auf Grund des Satzes in Nr. 59 $\frac{4a^2}{3\pi}$. Demnach hat man zur Bestimmung von μ_1 nach Nr. 58 die Gleichung:

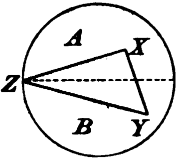


Fig. 11.

$$(\pi a^2)^2\mu_1 = \left(\frac{\pi a^2}{2}\right)^2\frac{11a^2}{18\pi} + \left(\frac{\pi a^2}{2}\right)^2\frac{11a^2}{18\pi} + 2\left(\frac{\pi a^2}{2}\right)^2\frac{4a^2}{3\pi}$$

und findet daraus

$$\mu_1 = \frac{35a^2}{36\pi}.$$

Die Bestimmung der mittleren Fläche μ eines beliebigen Dreiecks XYZ im Kreise kann nun mit Hilfe der in Nr. 63 entwickelten Differentialgleichung erfolgen; diese lautet im vorliegenden Falle:

$$\pi a^2 d\mu = 6\pi a da\left(\frac{35a^2}{36\pi} - \mu\right),$$

wenn man sich den Kreis durch einen Ring von der Breite da erweitert denkt; dann aber ist auch

$$\frac{\mu + d\mu}{\pi a^2 + 2\pi a da} = \frac{\mu}{\pi a^2} = \frac{d\mu}{2\pi a da};$$

mit Benutzung dieser Relation ergibt sich ohne Integration:

$$\mu = \frac{35a^2}{48\pi}.$$

Hiernach ist die Wahrscheinlichkeit eines nichtkonvexen Vierecks im Kreise $\frac{35}{12\pi^2}$, die eines konvexen $1 - \frac{35}{12\pi^2} = 0{,}7045 \cdots$.

67. Willkürliche Gerade in der Ebene. Der in Nr. 50 getroffenen Grundannahme entsprechend hat man sich die transfinite Menge der Geraden in der Ebene, wenn betreffs ihrer Erzeugung keine besondere Bestimmung getroffen ist, als homogen vorzustellen, d. h. bei gegebener Richtung sind alle Lagen, und bei gegebener Lage (Punkt) sind alle Richtungen gleichmäßig vertreten.

Die Menge der Geraden, die durch einen Punkt gehen und zwischen zwei Grenzrichtungen sich befinden, hat das Bogenmaß des Winkels dieser Grenzrichtungen zum Maße; insbesondere ist durch π die Menge aller Geraden durch einen Punkt gemessen.[1])

Die Menge paralleler Geraden zwischen zwei bestimmten Grenzlagen wird durch den Normalabstand dieser Grenzlage gemessen.

Die Menge der Geraden, deren Abstand von einem festen Punkte zwischen l und $l + dl$ liegt und deren Neigungswinkel (oder auch der Neigungswinkel des Lotes) zwischen θ und $\theta + d\theta$ variiert, hat zum Maße das Produkt $dl d\theta$.

Die Menge der Geraden, welche nach Lage und Richtung vorgeschriebenen Bedingungen genügen, wird durch das Integral $\iint dl\, d\theta$ bestimmt, wobei das Integrationsgebiet den Bedingungen entsprechend zu begrenzen ist.

Von besonderer Wichtigkeit ist es, ein Maß für die Menge der Geraden zu finden, welche eine geschlossene Kurve schneiden. Von der Erwägung ausgehend, daß durch alle Punkte der Ebene gleichviele Geraden hindurchgehen und daß die Menge der Punkte auf einer Kurve durch deren Länge gemessen wird, kommt man zu dem Schlusse, daß die Länge L der Kurve auch ein Maß für die Menge der sie schneidenden Geraden sei. Dieser Schluß ist aber nur dann statthaft, wenn jede Gerade gleich oft zur Zählung kommt, und dies ist der Fall, wenn die Kurve nach außen *durchaus konvex* ist, weil dann jede Gerade sie zweimal schneidet und daher zweimal gezählt wird.

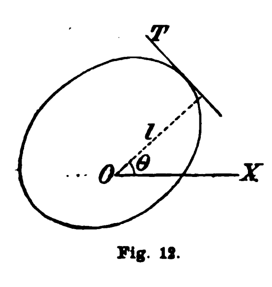


Fig. 12.

Nach den vorhin entwickelten Grundsätzen wäre die in Rede stehende Menge durch $\iint dl d\theta$ zu bestimmen. Integriert man bei innenliegendem Ursprung O, Fig. 12, und festem θ

1) Einer Geraden wird dabei nur *eine* Richtung zugeschrieben.

zunächst nach l, so ergibt sich, weil dann θ aller Werte zwischen 0 und 2π fähig ist, der Ausdruck

$$\int_0^{2\pi} l\,d\theta,$$

welcher tatsächlich die Länge der Kurve vorstellt.[1])

Ist die Kurve nicht durchweg konvex, so tritt an die Stelle ihrer Länge die Länge eines sie umspannenden Fadens; denn jede Gerade, welche die Umspannung schneidet, schneidet auch die Kurve, und umgekehrt.

Liegen in einer Ebene zwei konvexe Konturen derart, daß der eine (von der Länge L) den andern (von der Länge L') umschließt, so kommt die Wahrscheinlichkeit, daß eine den äußern Umriß schneidende, sonst beliebig gezogene Gerade auch den innern Umriß schneide, gleich dem Verhältnis $\frac{L'}{L}$ der Längen. Dieser Satz läßt sich bei der Lösung von Aufgaben über geometrische Wahrscheinlichkeit in mannigfacher Weise verwerten.

68. Beispiel XLI. *Eine Scheibe mit konvexem Umriß wird auf eine Ebene geworfen, die mit äquidistanten Parallelen überzogen ist; es ist die Wahrscheinlichkeit zu bestimmen, daß die Scheibe eine der Parallelen decke, vorausgesetzt, daß sie vermöge ihrer Form und Größe nicht mehrere zugleich decken kann.*

Diese Aufgabe kann in die folgende umgewandelt werden: Die Scheibe ist von einem Kreis umschlossen, dessen Durchmesser gleich ist dem Abstand $2a$ der Parallellinien; wie groß ist die Wahrscheinlichkeit, daß eine Gerade, welche den Kreis schneidet, auch den Umfang der Scheibe treffe? Denn wird die Scheibe in fester Verbindung mit dem Kreise auf die besagte Ebene geworfen, so *muß* eine der Parallelen den Kreis schneiden; es handelt sich also nur mehr um die Wahrscheinlichkeit, daß sie auch die Scheibe treffe, und diese ist $\frac{L}{2\pi a}$, wenn L den Umfang der Scheibe bedeutet.

Diese Lösung des Nadelproblems (Nr. 53) ist hierin als spezieller Fall enthalten; denn man kann die Nadel, deren Länge $2c$ ist, als Grenzform einer konvexen Scheibe vom Umfange $4c$ ansehen; demnach ist die Wahrscheinlichkeit, daß die auf die Ebene geworfene Nadel eine der Parallelen kreuze, $p = \frac{4c}{2\pi a} = \frac{2c}{\pi a}$.

Die vorstehende Aufgabe verdankt ihren Ursprung einem in Frankreich unter dem Namen „jeu du joint couvert" geübten Spiele, bei welchem eine Münze auf einen durch Fugen in gleichbreite

---

1) Cauchy, Compt. rend. XII (1841), p. 1060 ff.

Streifen zerlegten Fußboden geworfen und darauf gewettet wird, daß sie eine Fuge decken werde. Aufgaben ähnlicher Art sind von Buffon[1]) zuerst gestellt worden; Verallgemeinerungen derselben hat Barbier[2]) behandelt.

69. Beispiel XLII. *Die Wahrscheinlichkeit zu finden, daß eine Gerade, welche eine konvexe Kurve L schneidet, auch eine außerhalb ihr befindliche konvexe Kurve L schneiden werde.*

Die Lösung erfordert die Bestimmung des Maßes für die Menge jener Geraden, welche beide Kurven zugleich schneiden.

Zu diesem Ende verzeichne man die inneren und äußeren gemeinsamen Tangenten der Kurven und überlege an der Hand der Fig. 13 wie folgt: Die Menge der Geraden, welche den konvexen Umriß $OPLQO$, und derjenigen, welche den konvexen Umriß $OP'L'Q'O$ schneiden, ist durch die Summe der Umfänge beider, also durch die Länge X des beide Kurven umspannenden und sich kreuzenden Bandes gemessen. Unter diesen Geraden kommen diejenigen, welche L und L' zugleich schneiden, zweimal vor. Vergleicht man die beschriebene Menge mit der Mannigfaltigkeit jener Geraden, welche den konvexen Umriß $LRR'L'S'SL$ schneiden, die gemessen wird durch die Länge Y des beide Kurven umspannenden und sich nicht kreuzenden Bandes, so erkennt man leicht, daß in dieser Mannigfaltigkeit alle Geraden der erstbetrachteten Menge auch vorkommen, daß jedoch die L und L' zugleich schneidenden darin nur einfach auftreten; daher mißt die Differenz

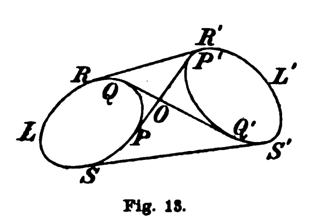


Fig. 13.

$$X - Y$$

die Menge dieser letzteren Geraden.

Da L (darunter die Länge der Kurve verstanden) die Menge der Geraden bedeutet, welche die gleichnamige Kurve schneiden, so ist die Wahrscheinlichkeit, daß eine willkürliche unter diesen Geraden auch L' treffe,

$$p = \frac{X - Y}{L}.$$

Ein einfaches Beispiel hierzu bietet die folgende Aufgabe: Die Wahrscheinlichkeit zu bestimmen, daß eine Gerade, welche die Halbkreislinie ACB (Fig. 14) trifft, auch die supplementäre Halbkreislinie ADB schneide.

Zunächst denke man sich jede der Halbkreislinien durch den Durchmesser AOB zu einem konvexen Kontur ergänzt. Man erkennt nun leicht, daß

1) Vgl. hierzu Nr. 48. 2) Journal Liouville, (2) V, 1860.

$$L = \pi a + 2a, \qquad X = 2\pi a + 4a, \qquad Y = 2\pi a$$

ist; infolgedessen hat man

$$p = \frac{4a}{\pi a + 2a} = \frac{4}{\pi + 2} = 0{,}7779 \cdots .$$

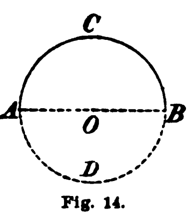


Fig. 14.

Indessen kann die Aufgabe auch mittels des Schlußsatzes von Nr. 67 gelöst werden. Die Gerade, die ACB schneidet, trifft auch und nur dann den supplementären Halbkreis, wenn sie den Durchmesser AB schneidet; dieser aber kann als die Grenzform eines geschlossenen konvexen Umrisses von der Länge $4a$ innerhalb des ebenfalls konvexen Umrisses $ACBOA$ von der Länge $\pi a + 2a$ betrachtet werden.

70. Beispiel XLIII. *Es ist die Wahrscheinlichkeit zu bestimmen, daß zwei beliebige Sekanten eines konvexen geschlossenen Umrisses einander innerhalb desselben schneiden.*

Die Länge des Umrisses sei L, die von ihm eingeschlossene Fläche heiße F.

Die auf der erstgezogenen Sekante ausgeschnittene Sehne c kann als die Grenzform eines geschlossenen konvexen Umrisses von der Länge $2c$ angesehen werden, der innerhalb des gegebenen liegt; daher ist $\frac{2c}{L}$ die Wahrscheinlichkeit, daß die zweite Sekante *diese* Sehne, also die erste Sekante innerhalb L schneide. Sind l, θ die Parameter der ersten Sekante in der in Nr. 67 erklärten Weise, so ist $\frac{dl d\theta}{L}$ die Wahrscheinlichkeit der Sehnenlänge c, folglich

$$\frac{2c}{L^2} dl d\theta$$

die zusammengesetzte Wahrscheinlichkeit, daß bei dieser Lage der ersten Sekante ein innerer Schnitt erfolge. Die totale Wahrscheinlichkeit eines solchen Schnittes ist also:

$$p = \frac{2}{L^2} \int\!\!\int c\, dl\, d\theta = \frac{2}{L^2} \int d\theta \int c\, dl;$$

nun bestimmt $\int c\, dl$, auf alle Sekanten einer bestimmten Richtung bezogen, die Fläche F, und die noch übrige Integration nach θ ist zwischen den Grenzen 0 und π vorzunehmen; daher ist endgültig

$$p = \frac{2\pi F}{L^2}.$$

Zu einem andern Ausdruck für p gelangt man, wenn man die

mittlere Länge der Sehne c, $\mathfrak{D}(c)$ zu Hilfe nimmt; dann ist nämlich unmittelbar (Nr. 56)

$$p = \frac{2\,\mathfrak{D}(c)}{L}.$$

Durch Vergleichung beider Resultate ergibt sich die mittlere Sehnenlänge eines konvexen Konturs von der Länge L und der umschlossenen Fläche F:

$$\mathfrak{D}(c) = \frac{\pi F}{L}.$$

Es ist wohl zu beachten, daß hier die Sehne als der auf einer beliebigen Geraden der Ebene durch die Kurve gebildete Abschnitt verstanden wird (Nr. 50).

Aus diesen allgemeinen Formeln ergibt sich beispielsweise, daß zwei beliebige Sekanten eines Quadrates von der Seite a sich mit der Wahrscheinlichkeit $\frac{\pi}{8} = 0{,}3927 \cdots$ innerhalb des Quadrates schneiden, und daß die mittlere Länge der auf den Sekanten ausgeschnittenen Sehnen gleichkommt $\frac{\pi a}{4}$; daß zwei beliebige Sekanten eines Kreises vom Halbmesser a sich mit der Wahrscheinlichkeit $\frac{1}{2} = 0{,}5$ innen schneiden und die mittlere Länge der Sehne $\frac{\pi a}{2}$ beträgt.

71. Das Bertrandsche Paradoxon. Unter dieser vermutlich von Poincaré[1]) stammenden Bezeichnung ist ein von Bertrand[2]) aufgestelltes Problem Gegenstand wiederholter Kritik und Untersuchung geworden. Die Fragestellung, wie sie der Urheber formulierte, lautet: „In einem Kreise wird eine Sehne *beliebig* gezogen. Welches ist die Wahrscheinlichkeit, daß sie größer ist als die Seite des eingeschriebenen gleichseitigen Dreiecks?“

Der Ursprung der Paradoxie, die man glaubte in diesem Problem gefunden zu haben, liegt darin, daß man den Begriff einer „beliebigen Sehne“ verschieden deuten, sie verschiedenen transfiniten Mannigfaltigkeiten zuzählen kann, je nach der Art, wie man die Sehne erzeugt denkt (Nr. 50). Bertrand selbst hat hierüber drei verschiedene Annahmen gemacht und ist jedesmal, was gar nichts Überraschendes an sich hat, zu einem andern Resultat gekommen; er faßt schließlich seine Anschauung über die Sache in die Worte: „Welche unter diesen Antworten ist die wahre? Keine von den drei ist falsch, keine ist zutreffend, die Frage ist schlecht gestellt.“

Von diesen drei Behauptungen ist die mittlere unrichtig; jede der Lösungen ist unter der ihr zugrundeliegenden Annahme wahr.

1) Calc. d. probab., p. 94.
2) Calc. d. probab., p. 4.

Was die Fragestellung betrifft, so verhält es sich mit ihr ebenso wie bei allen Problemen der geometrischen Wahrscheinlichkeit, bei denen die „beliebige Annahme“ dieses oder jenes Elements verschiedener Auslegungen fähig ist.

Bertrands Annahmen sind die folgenden: 1. Es wird ein Endpunkt der Sehne angenommen, diese dann in beliebiger Richtung gezogen. 2. Es wird die Richtung der Sehne angenommen, diese dann durch einen beliebigen Punkt des zur Richtung normalen Durchmessers geführt. 3. Es wird der Mittelpunkt der Sehne beliebig gewählt.

Man kann diese Annahmen vermehren und beispielsweise auch so vorgehen: 4. Nachdem ein Endpunkt angenommen, wird die Sehne durch einen beliebigen Punkt der Kreisfläche gezogen. 5. Es werden beide Endpunkte beliebig angenommen. 6. Die Sehne wird durch zwei in der Kreisfläche beliebig gewählte Punkte geführt usw.

In allen Fällen wird man aber schließlich an der Grundannahme festhalten, das Gebiet, auf das sich die Willkür bezieht, als homogen vorauszusetzen.

Wollte man der Mehrdeutigkeit dadurch ausweichen, daß man jener Auffassung vor allen andern den Vorzug gibt, welche der Natur der Sache, oder besser gesagt unserer Gewohnheit des Sehnenziehens, am besten entspricht, so wäre auch eine Entscheidung hierüber kaum zu treffen; am wenigsten entspricht jedenfalls die Modalität (3).

Die Lösung des Problems gestaltet sich unter den aufgezählten Annahmen wie folgt.

1. Ist einmal der eine Endpunkt angenommen, so denke man sich von ihm aus die beiden Sehnen von der Länge der Dreiecksseite gezogen; fällt die beliebig angenommene Sehne zwischen diese, so ist sie größer als die Dreiecksseite. Da nun das Maß der Gesamtheit der Richtungen durch den Punkt π, das Maß der Gesamtheit der günstigen $\frac{\pi}{3}$ ist, so ist $p = \frac{1}{3}$.

2. Diese Modalität entspricht der Auffassung der Sehne als Ausschnitt auf einer in der Ebene des Kreises gezogenen Geraden, die den Kreis schneidet; schneidet sie auch den konzentrischen Kreis vom halben Radius, so ist die Sehne, und nur dann, länger als die Dreiecksseite. Da die Umfänge der Kreise (den Radius als Einheit angenommen) 2π und π sind, so ist (Nr. 67) $p = \frac{1}{2}$.

3. Fällt der in der Kreisfläche angenommene Punkt in die Fläche des konzentrischen Kreises vom halben Radius, so wird die durch ihn gezogene kürzeste Sehne länger als die Dreiecksseite. Da die Punktmengen der beiden Kreisflächen die Inhalte π und $\frac{\pi}{4}$ besitzen, so ist $p = \frac{1}{4}$.

4. Nachdem der Endpunkt angenommen, ziehe man von ihm aus die beiden Sehnen von der Länge der Dreiecksseite. Fällt dann der beliebig angenommene Punkt in den zwischen den Sehnen enthaltenen Flächenteil, so tritt ein günstiger Fall ein. Der Menge π der möglichen Punkte steht die Menge $\frac{\pi}{3} + \frac{\sqrt{3}}{2}$ der günstigen gegenüber; also ist $p = \frac{1}{3} + \frac{\sqrt{3}}{2\pi} = 0{\cdot}609\cdots$.

5. Diese Auffassung fällt mit der unter 1. zusammen, weil der gleichmäßigen Verteilung der Punkte auf dem Umfang auch eine gleichförmige Verteilung der vom angenommenen ersten Endpunkt nach ihnen führenden Richtungen entspricht; demnach ist auch hier $p = \frac{1}{3}$.

6. Durch den Kreis vom Radius $OR = \frac{1}{2}$, Fig. 15, zerfällt der gegebene Kreis in zwei Gebiete. Fällt der erstangenommene Punkt M auf das innere Gebiet, wofür die Wahrscheinlichkeit $\frac{1}{4}$ besteht, so ist die Sehne sicher, wie auch der zweite Punkt N angenommen werden möge, länger als die Dreiecksseite. Fällt der erste Punkt M in das äußere Gebiet auf einen Kreisring mit den Radien $OM = x$ und $x + dx$, wofür die Wahrscheinlichkeit $\frac{2\pi x\,dx}{\pi} = 2x\,dx$ besteht,

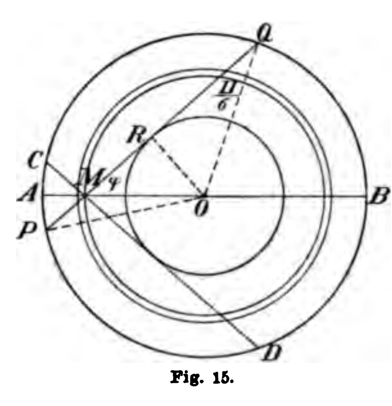


Fig. 15.

so ergibt sich ein günstiger Fall nur dann, wenn der zweite Punkt N in eine der Flächen $MQBD$, $MPAC$ zu liegen kommt. Setzt man also $MQBD + MPAC = X$, so ist

$$p = \frac{1}{4} + \frac{2}{\pi}\int_{\frac{1}{2}}^{1} X x\,dx.$$

Nun ist

$$MQB = OQB + \Delta MOQ = OQB + 2\Delta MOR + \Delta MOP$$

$$MPA = OPA - \Delta MOP,$$

daher

$$X = 2\,(OQB + OPA + 2\Delta MOR);$$

mit Benützung des Winkels $OMQ = \varphi$ und unter Beachtung, daß die Winkel bei P und Q je 30^0 betragen, ergibt sich

$$X = 2\varphi + x \cos\varphi,$$

$$x = \frac{1}{2 \sin\varphi}, \quad dx = -\frac{\cos\varphi\, d\varphi}{2 \sin^2\varphi}$$

und hiermit

$$p = \frac{1}{4} + \frac{1}{\pi}\int\limits_{\frac{\pi}{6}}^{\frac{\pi}{2}} \left(2\varphi + \frac{\cos\varphi}{2\sin\varphi}\right) \frac{\cos\varphi\, d\varphi}{2\sin^2\varphi};$$

es ist aber

$$\int \frac{\varphi \cos\varphi\, d\varphi}{\sin^2\varphi} = -\frac{\varphi}{2\sin^2\varphi} - \frac{\operatorname{cotg}\varphi}{2},$$

$$\int \frac{\cos^2\varphi\, d\varphi}{4 \sin^4\varphi} = -\frac{\operatorname{cotg}^3\varphi}{12},$$

daher schließlich[1])

$$p = \frac{4\pi + 9\sqrt{3}}{12\pi} = 0{\cdot}746\cdots.$$

II. Abschnitt. Wahrscheinlichkeiten, betreffend die Ergebnisse wiederholter Beobachtungen.

§ 1. Das Theorem von Bernoulli.

72. Entwicklung der Fragestellung. Aus der Verwirklichung gewisser allgemeiner Bedingungen möge mit Notwendigkeit eines der Ereignisse E, $\bar{E}$ hervorgehen; für das Eintreffen des ersten bestehe die Wahrscheinlichkeit p, für das zweite die Wahrscheinlichkeit q, so daß $p + q = 1$ ist. Die Verwirklichung (Versuch, Beobachtung) soll nicht einmal, sondern s-mal geschehen; jedoch sei ausdrücklich vorausgesetzt, daß der Komplex der allgemeinen Bedingungen in dem Sinne unveränderlich bleibe, daß die Wahrscheinlichkeiten p, q bei jedem Versuche dieselben Werte besitzen. Das einfachste Schema für diesen Sachverhalt kann in wiederholten Ziehungen einer Kugel aus einer Urne erblickt werden, welche weiße und schwarze Kugeln in bestimmtem Mengenverhältnis enthält; die gezogene Kugel wird jedes-

1) Man vergleiche zu diesem Problem: H. Poincaré, Calc. d. probab., p. 94—96 und 98; R. de Montessus et G. Lechalas, Un paradoxe du calc. d. probab. Nouv. Ann. IV (3), 1903, p. 21—31, 343—348; E. Cesàro, Considerazioni sul concetto di probab., Periodico di matemat. VI (1891), p. 1—13, 49—62. — R. Lämmels Dissertation (p. 29—32, 38—39) ist in den bezüglichen Ausführungen unklar und widersprechend; das dort für die Modalität 6 angegebene Resultat $p = \frac{5\pi + 9\sqrt{3}}{16\pi} = 0{\cdot}622\cdots$ ist unrichtig.

mal zurückgelegt und unter die andern gemengt, ehe der nächste Zug geschieht; das Erscheinen einer weißen Kugel gelte als Ereignis E, das einer schwarzen als Ereignis $\bar{E}$.

Wenn man bloß die Wiederholungszahlen der Ereignisse E, $\bar{E}$ in Betracht zieht und von ihrer Reihenfolge, falls sie überhaupt feststellbar wäre, absieht, so sind $s+1$ verschiedene Gesamterfolge der s Beobachtungen möglich. Es kann

das Ereignis E s-mal, das Ereignis $\bar{E}$ gar nicht,
" " " $s-1$ " " " " 1 mal
" " " $s-2$ " " " " 2 "
. . . .
das Ereignis E gar nicht, das Ereignis $\bar{E}$ s-mal

eintreten. Die Wahrscheinlichkeiten dieser Kombinationen sind durch die aufeinander folgenden Glieder der nach fallenden Potenzen von p geordneten Entwicklung von

$$(p+q)^s$$

bestimmt (s. Nr. 39); insbesondere stellt das Glied

$$\frac{s!}{m!\,n!}\,p^m q^n \qquad (m+n=s) \qquad (1)$$

die Wahrscheinlichkeit $\mathfrak{W}(E^m \bar{E}^n)$ dar, daß der Gesamterfolg in der m-maligen Wiederholung von E und der n-maligen Wiederholung $\bar{E}$ bestehen werde oder, falls er schon vorliegt, bestehe.

Es mag hervorgehoben werden, daß diese Wahrscheinlichkeitsmessung für einen Gesamterfolg geradeso wie die Wahrscheinlichkeitsbestimmung für den einzelnen Versuch lediglich auf der Wahrscheinlichkeitsdefinition beruht und demgemäß folgenden Sinn hat: In dem Gesamtumfang der Ausführungsmöglichkeiten der s Versuche machen diejenigen, welche mit Notwendigkeit zu dem in Rede stehenden Gesamterfolg hinführen, einen Teil aus, dessen Verhältnis zum Ganzen durch die Zahl (1) bestimmt ist.

Hinsichtlich der möglichen Erfolge der s Versuche sind nun zwei Fragen von besonderer Bedeutung: Welcher unter den Erfolgen ist der wahrscheinlichste? Welche Wahrscheinlichkeit besteht dafür, daß sich ein Erfolg einstellen werde, der von dem wahrscheinlichsten nicht mehr als innerhalb vorgegebener Grenzen abweicht?

Der Sinn des Wortes „Abweichung" ist hier der folgende: Sind m', n' die Wiederholungszahlen von E und $\bar{E}$ in der wahrscheinlichsten Kombination, so weist eine Kombination mit den Wiederholungszahlen m, n die Abweichung $m-m' = -(n-n')$ auf; um so viel übertrifft (wenn $m-m'>0$) die Wiederholungszahl von E in der

letzteren Kombination jene in der wahrscheinlichsten Kombination, oder um so viel bleibt sie hinter ihr zurück (wenn $m - m' < 0$).

Die Beantwortung der aufgestellten Fragen bietet keine prinzipiellen Schwierigkeiten dar. Die erste erfordert die Feststellung des größten Gliedes in der Entwicklung von $(p+q)^s$; die Exponenten von p, q in diesem Gliede kennzeichnen die wahrscheinlichste Kombination, der Wert des Gliedes selbst gibt ihre Wahrscheinlichkeit. Um die zweite Frage zu beantworten, hat man jene Glieder der Entwicklung auszurechnen, welche Abweichungen innerhalb der vorgezeichneten Grenzen aufweisen; ihre Summe gibt die Wahrscheinlichkeit, daß diese Grenzen nicht überschritten werden.

Ist s eine nur mäßige Zahl und sind auch p, q durch einfache, d. h. mit wenigen Ziffern darstellbare Zahlen ausgedrückt, so verursacht die Ausführung dieser Rechnungen keine große Mühe. Wir wollen zwei Beispiele in dieser Art erledigen, um daran einige Wahrnehmungen zu knüpfen.

Aus einer Urne, welche zwei weiße und eine schwarze Kugel enthält, mögen 6 Ziehungen vorgenommen werden.

Die Entwicklung

$$\left(\frac{2}{3}+\frac{1}{3}\right)^6=\left(\frac{2}{3}\right)^6+6\left(\frac{2}{3}\right)^5\frac{1}{3}+15\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2+20\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3+15\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4+6\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5+\left(\frac{1}{3}\right)^6$$
$$=\frac{64}{729}+\frac{192}{729}+\frac{240}{729}+\frac{160}{729}+\frac{60}{729}+\frac{12}{729}+\frac{1}{729}$$

zeigt, daß die Kombination, welche 4 weiße und 2 schwarze Kugeln umfaßt, unter allen die wahrscheinlichste ist; auf die Frage nach der Wahrscheinlichkeit, daß die Wiederholungszahl der weißen Kugel nicht über 5 und nicht unter 3 falle, die Abweichung von der wahrscheinlichsten Wiederholungszahl also nicht mehr als 1 nach auf- oder abwärts betrage, gibt sie die Antwort:

$$\frac{192}{729}+\frac{240}{729}+\frac{160}{729}=\frac{592}{729}.$$

Aus derselben Urne sollen doppelt so viel, das sind 12 Ziehungen vorgenommen werden.

Die 13 Glieder der nach fallenden Potenzen von $\frac{2}{3}$ geordneten Entwicklung von $\left(\frac{2}{3}+\frac{1}{3}\right)^{12}$ sind Brüche mit dem Nenner 531441 und den Zählern:

4096, 24576, 67584, 112640, 126720, 101376, 59136, 25344, 7920, 1760, 264, 24, 1.

Daraus ist zu entnehmen, daß die Kombination von 8 weißen und

4 schwarzen Kugeln die wahrscheinlichste ist; daß ferner die Wahrscheinlichkeit, weiß werde nicht öfter als 10mal und nicht seltener als 6mal sich einstellen, gleich ist

$$\frac{67584}{531441} + \frac{112640}{531441} + \frac{126720}{531441} + \frac{101376}{531441} + \frac{59136}{531441} = \frac{467456}{531441}.$$

An diesen Beispielen sei folgendes hervorgehoben.

Bei 6 Versuchen hat die wahrscheinlichste Kombination die Wahrscheinlichkeit:

$$\frac{240}{729} = 0{,}3030 \cdots,$$

bei 12 Versuchen — die übrigen Umstände blieben unverändert — die Wahrscheinlichkeit:

$$\frac{126720}{531441} = 0{,}2380 \cdots;$$

diese Wahrscheinlichkeit hat sich mit der wachsenden Anzahl der Versuche vermindert. Diese Erscheinung kann nicht überraschen, wenn man beachtet, daß mit der Anzahl der Versuche auch die Menge der möglichen Erfolge wächst.

Für die Grenzen 5, 3 der Wiederholungszahl weißer Kugeln ergab sich bei 6 Versuchen die Wahrscheinlichkeit:

$$\frac{529}{729} = 0{,}7474 \cdots;$$

für die Grenzen 10, 6 bei 12 Versuchen die Wahrscheinlichkeit:

$$\frac{467456}{531441} = 0{,}8796 \cdots.$$

Das Verhältnis der Grenzen zur Gesamtzahl der Versuche ist beidemal dasselbe, nämlich $\frac{5}{6} = \frac{10}{12}$ und $\frac{3}{6} = \frac{6}{12}$, die Wahrscheinlichkeit der Nichtüberschreitung dieser Grenzen fiel das zweite Mal größer aus.

Die Reihe der Wahrscheinlichkeiten ist asymmetrisch, nach der Seite des Ereignisses mit der größeren Wahrscheinlichkeit kürzer und aus größeren Zahlen bestehend. Sie würde symmetrisch ausfallen, wenn $p = q = \frac{1}{2}$ wäre.

Bei *sehr großen* Werten von s und weiteren Grenzen der Abweichung wird die direkte Erledigung der gestellten Fragen wegen der Weitläufigkeit und Umständlichkeit der erforderlichen Rechnungen physisch undurchführbar. Für solche Fälle sind seit der Zeit, da Jakob Bernoulli diese Untersuchungen in Angriff genommen[1]),

1) Ars conjectandi, pars quarta, p. 210—239 (Ostwalds Klassiker Nr. 108, p. 71—107).

Näherungsmethoden ausgebildet worden, um welche sich insbesondere A. de Moivre[1]), J. Stirling[2]), C. Maclaurin[3]), L. Euler[4]) und P. S. Laplace[5]) verdient gemacht haben. Mit der Entwicklung dieser Analyse und ihrer Anwendung auf verschiedene Probleme beschäftigen sich die folgenden Nummern.

73. Wahrscheinlichstes Ergebnis einer Versuchsreihe. Seine Ermittlung ist gleichbedeutend mit der Aufsuchung des größten Gliedes der Entwicklung von $(p+q)^s$. Ist $\frac{s!}{m!\,n!}p^m q^n$ dieses Glied, so steht es mit dem vorangehenden und nachfolgenden in der Beziehung:

$$\frac{s!}{(m+1)!\,(n-1)!}p^{m+1}q^{n-1} < \frac{s!}{m!\,n!}p^m q^n > \frac{s!}{(m-1)!\,(n+1)!}p^{m-1}q^{n+1},$$

woraus einerseits

$$\frac{m+1}{n}\frac{q}{p} > 1,$$

andererseits

$$\frac{n+1}{m}\frac{p}{q} > 1$$

folgt; eliminiert man hieraus n und q mittels der Gleichungen $m+n=s$ und $p+q=1$, so ergeben sich für m die Grenzen:

$$sp - q < m < sp + p. \tag{1}$$

Da das Intervall dieser Grenzen 1 beträgt, so ist durch sie, sofern sie nicht ganze Zahlen sind, m als ganze Zahl *eindeutig* bestimmt. Ist hingegen $sp-q$ und infolgedessen auch $sp+p$ eine ganze Zahl, so ergeben sich zwei um eine Einheit verschiedene Werte von m, die Entwicklung hat *zwei* gleiche Glieder, welche größer sind als alle übrigen.

Wie dem auch sei, der Quotient $\frac{m}{s}$ unterscheidet sich höchstens um $\frac{1}{s}$ von p und ebenso der Quotient $\frac{n}{s}$ höchstens um $\frac{1}{s}$ von q; ist sp eine ganze Zahl, dann ist $m=sp$ die der Relation (1) entsprechende Zahl und $\frac{m}{s}$ dann genau gleich p, ebenso $\frac{n}{s}$ genau gleich q.

Unter allen Ergebnissen, die in s Versuchen möglich sind, ist hiernach dasjenige am wahrscheinlichsten, in welchem das Verhältnis $m:n$ der Wiederholungszahlen von E und $\bar{E}$ dem Verhältnis $p:q$ der Wahrscheinlichkeiten gleich oder am nächsten ist.

---

1) Doctrine of chances (3. Aufl. 1756), p. 243—254; Miscellanea analytica (1730), Supplem.

2) Methodus differentialis (1730). 3) Treatise of Fluxions (1742).

4) Institutiones calculi differentialis (1755).

5) Théorie analyt. d. probab. (1812), chap. III.

Die weiter folgende Analyse läßt es zu,

$$m = sp, \quad n = sq$$

zu setzen und daher das größte Glied in der Form

$$T_0 = \frac{s!}{(sp)!\,(sq)!} p^{sp} q^{sq} \tag{2}$$

auch dann zu schreiben, wenn sp, sq nicht ganze Zahlen sein sollten.

Zur Erläuterung dienen die folgenden Zahlenbeispiele. Aus einer Urne mit 3 weißen und 2 schwarzen Kugeln werden 400, beziehungsweise 623, 624 Ziehungen gemacht; welches ist in jedem dieser drei Fälle das wahrscheinlichste Resultat?

Die Wahrscheinlichkeit einer weißen Kugel ist $p = \frac{3}{5}$, die einer schwarzen $q = \frac{2}{5}$.

Bei $s = 400$ ist $sp = 240$, $sq = 160$, und da dies ganze Zahlen sind, so besteht die wahrscheinlichste Kombination aus 240 weißen und 160 schwarzen Kugeln.

In dem Falle $s = 623$ sind $373\frac{2}{5}$, $374\frac{2}{5}$ die Grenzen des m einschließenden Intervalls; daher ist $m = 374$, $n = 249$.

In dem Falle $s = 624$ findet man 374, 375 als Grenzen des Intervalls für m; man darf daher ebensowohl $m = 374$, $n = 250$ wie $m = 375$, $n = 249$ setzen.

74. Näherungsweise Darstellung des größten Gliedes. Ist s eine große Zahl, so darf man die in dem maximalen Gliede T_0 auftretenden Faktoriellen mittels der Stirlingschen Formel ausdrücken; darnach ist

$$s! = s^{s+\frac{1}{2}} e^{-s} \sqrt{2\pi}$$

$$(sp)! = (sp)^{sp+\frac{1}{2}} e^{-sp} \sqrt{2\pi}$$

$$(sq)! = (sq)^{sq+\frac{1}{2}} e^{-sq} \sqrt{2\pi};$$

daraus ergibt sich für den Binomialkoeffizienten $\frac{s!}{(sp)!\,(sq)!}$ der angenäherte Wert

$$\frac{1}{p^{sp+\frac{1}{2}} q^{sq+\frac{1}{2}} \sqrt{2\pi s}}$$

und für das größte Glied der angenäherte Wert:

$$T_0 = \frac{1}{\sqrt{2\pi spq}}. \tag{3}$$

Bei demselben p, q nimmt also die Wahrscheinlichkeit der wahrscheinlichsten Kombination mit wachsender Versuchszahl, und zwar im Verhältnis ihrer reziproken Quadratwurzel, ab. Bei gegebenem s hat T_0 den kleinsten Wert, wenn $p = q = \frac{1}{2}$ ist.

Daß bei 1000 Würfen mit einer Münze gleich oft Wappen und Schrift sich einstelle, hat die Wahrscheinlichkeit

$$\frac{1}{\sqrt{500\pi}} = 0{,}025231\,;$$

daß in 1000 Ziehungen aus einer Urne mit 4 weißen und 1 schwarzen Kugel 800 weiße und 200 schwarze Kugeln erscheinen, ist mit der Wahrscheinlichkeit

$$\frac{1}{\sqrt{320\pi}} = 0{,}031539$$

zu erwarten.

Um zu zeigen, daß die Formel (3) auch schon bei mäßigem s ziemliche Annäherung gibt, benützen wir die beiden Beispiele in Nr. 72. Dort war $p = \frac{2}{3}$, $q = \frac{1}{3}$

und für $s = 6$ der strenge Wert von $T_0 = 0{,}3030$,
der angenäherte ist $0{,}3455$;
für $s = 12$ der strenge Wert von $T_0 = 0{,}2380$,
der angenäherte ist $0{,}2443$.

Welche Näherung sie bei großem s liefert, soll an dem Falle $s = 1200$ durch Nebeneinanderstellung der strengen und der Näherungsrechnung dargetan werden; die linksstehende Rechnung ist mit Hilfe der in der Fußnote zu Nr. 11 an zweiter Stelle erwähnten Logarithmentafel der Fakultäten geführt.

$$T_0 = \frac{1200!}{800!\,400!}\,\frac{2^{800}}{3^{1200}}\,;$$

$$\begin{aligned}
\log 1200! &= 3175{,}802827\\
\log 800! &= 1976{,}887084\\
\log 400! &= 868{,}806414\\
\hline
&\ 330{,}109329\\
800 \log 2 &= 240{,}823997\\
\hline
&\ 570{,}933326\\
1200 \log 3 &= 572{,}545506\\
\hline
\log T_0 &= 0{,}387820-2\\
T_0 &= 0{,}024424
\end{aligned}$$

$$T_0 = \frac{3}{\sqrt{4800\pi}}$$

$$\begin{aligned}
\log 4800 &= 3{,}681241\\
\log \pi &= 0{,}497150\\
\hline
&\ 4{,}178391\\
&\ 2{,}089196\\
\log 3 &= 0{,}477121\\
\hline
\log T_0 &= 0{,}387925-2\\
T_0 &= 0{,}024430
\end{aligned}$$

75. Näherungsweise Darstellung eines Gliedes, das von dem maximalen eine Abweichung von der Ordnung $\sqrt{s}$ aufweist. Dasjenige Glied, welches dem größten Gliede vorangehend von diesem durch $l-1$ Glieder getrennt ist, hat den Ausdruck:

$$T_l = \frac{s!}{(sp+l)!\,(sq-l)!}\,p^{sp+l}q^{sq-l}; \tag{4}$$

das hierzu symmetrisch angeordnete Glied lautet:

$$T_{-l} = \frac{s!}{(sp-l)!\,(sq+l)!}\,p^{sp-l}q^{sq+l}. \tag{5}$$

Um für das erstgedachte Glied unter der Voraussetzung, daß l eine Zahl von der Größenordnung $\sqrt{s}$ sei, und unter Vernachlässigung von Größen von der Ordnung $\frac{1}{s}$ aufwärts einen Näherungswert zu erhalten, drücke man zunächst die Fakultäten mit Hilfe der Stirlingschen Formel aus; dabei benütze man die Umformung

$$(sp+l)! = (sp+l)^{sp+l+\frac{1}{2}}e^{-sp-l}\sqrt{2\pi} = (sp)^{sp+l+\frac{1}{2}}\left(1+\frac{l}{sp}\right)^{sp+l+\frac{1}{2}}e^{-sp-l}\sqrt{2\pi}$$

und ähnlich bei $(sq-l)!$. Dadurch ergibt sich für T_l der Ausdruck:

$$T_l = \frac{1}{\sqrt{2\pi spq}}\left(1+\frac{l}{sp}\right)^{-sp-l-\frac{1}{2}}\left(1-\frac{l}{sq}\right)^{-sq+l-\frac{1}{2}}.$$

Nun ist mit der oben angegebenen Approximation der natürliche Logarithmus

$$\begin{aligned}
\operatorname{Log}\left(1+\frac{l}{sp}\right)^{-sp-l-\frac{1}{2}} &= -\left(sp+l+\frac{1}{2}\right)\left(\frac{l}{sp}-\frac{l^2}{2s^2p^2}+\cdots\right)\\
&= -l-\frac{l}{2sp}-\frac{l^2}{2sp}+\frac{l^3}{6s^2p^2}+\cdots;\\
\operatorname{Log}\left(1-\frac{l}{sq}\right)^{-sq+l-\frac{1}{2}} &= -\left(sq-l+\frac{1}{2}\right)\left(-\frac{l}{sq}-\frac{l^2}{2s^2q^2}-\cdots\right)\\
&= \quad l+\frac{l}{2sq}-\frac{l^2}{2sq}-\frac{l^3}{6s^2q^2}-\cdots;
\end{aligned}$$

die Summe dieser Logarithmen reduziert sich auf

$$\frac{(p-q)l}{2spq}-\frac{l^2}{2spq}+\frac{l^3}{6s^2p^2}-\frac{l^3}{6s^2q^2}+\cdots,$$

das Produkt der beiden Klammerfaktoren von T_l kann also durch

$$e^{\frac{(p-q)l}{2spq} - \frac{l^2}{2spq} + \frac{l^3}{6s^2p^2} - \frac{l^3}{6s^2q^2} + \cdots}$$

oder durch

$$e^{-\frac{l^2}{2spq}}\left(1 + \frac{(p-q)l}{2spq} + \frac{l^3}{6s^2p^2} - \frac{l^3}{6s^2q^2} + \cdots\right)$$

ersetzt werden.

Hiernach ist

$$T_l = \frac{e^{-\frac{l^2}{2spq}}}{\sqrt{2\pi spq}}\left(1 + \frac{(p-q)l}{2spq} + \frac{l^3}{6s^2p^2} - \frac{l^3}{6s^2q^2} + \cdots\right);$$

daraus ergibt sich ohne weitere Rechnung:

$$T_{-l} = \frac{e^{-\frac{l^2}{2spq}}}{\sqrt{2\pi spq}}\left(1 - \frac{(p-q)l}{2spq} - \frac{l^3}{6s^2p^2} + \frac{l^3}{6s^2q^2} + \cdots\right).$$

Der gemeinsame Faktor beider Ausdrücke ist eine gerade Funktion von l; der zweite Faktor führt die Asymmetrie herbei; ist $p > q$, so ist $T_l > T_{-l}$.

Wird die Summe der beiden Wahrscheinlichkeiten T_l, T_{-l} genommen, so ergibt sich die Wahrscheinlichkeit, daß die Abweichung nach auf- oder abwärts l betrage, und diese ist

$$T_l + T_{-l} = \frac{2}{\sqrt{2\pi spq}}\, e^{-\frac{l^2}{2spq}}.$$

Sofern man also Glieder der Entwicklung zusammenfaßt, welche in bezug auf das größte Glied symmetrisch angeordnet sind, darf man innerhalb der angegebenen Grenzen von l und mit der bezeichneten Genauigkeit die Werte von T_l durch die *gerade* Funktion

$$T_x = \frac{1}{\sqrt{2\pi spq}}\, e^{-\frac{x^2}{2spq}} \tag{6}$$

darstellen.

Indessen liegt es in der Natur dieser Funktion, in ihrer außerordentlich raschen Abnahme mit wachsendem x, daß man sie auch über die angegebenen Grenzen hinaus anwenden darf, ohne einen erheblichen Einfluß auf die ziffermäßigen Resultate befürchten zu müssen. Es werden für beträchtliche Abweichungen sowohl die strengen wie auch die nach dem Gesetze (6) gerechneten Näherungswerte von T_l so außerordentlich klein, daß sie bei Lösung einer praktischen Frage nicht in Betracht kommen.

Die rasche Abnahme von T_l mag aus dem folgenden Beispiele

ersehen werden. Für $p = \frac{3}{5}$, $q = \frac{2}{5}$, $s = 2500$, in welchem Falle l von der Ordnung von 50, also ein mäßiges Vielfache von 50 sein darf, ergibt die Rechnung nach (6):

$$T_0 = \frac{1}{\sqrt{1200\pi}} = 0{,}016287,$$

$$T_{50} = T_0 e^{-\frac{25}{12}} = 0{,}0020280,$$

$$T_{80} = T_0 e^{-\frac{16}{3}} = 0{,}000078632,$$

$$T_{100} = T_0 e^{-\frac{25}{3}} = 0{,}0000039149,$$

$$T_{150} = T_0 e^{-\frac{75}{4}} = 0{,}00000000019318.$$

Die weitere Abnahme erfolgt so rasch, daß T_{500} erst an der 93. Stelle eine bedeutsame Ziffer hat.

76. Wahrscheinlichkeit, daß die Abweichung innerhalb gegebener Grenzen verbleibe. Die Wahrscheinlichkeit P, daß die Wiederholungszahl von E zwischen die Grenzen $sp - l$ und $sp + l$ und entsprechend dem die Wiederholungszahl von $\bar{E}$ zwischen $sq - l$ und $sq + l$ falle, mit andern Worten: daß die Abweichung des Erfolges von der wahrscheinlichsten Kombination dem Betrage nach l nicht überschreite, ist dargestellt durch die Summe

$$T_l + T_{l-1} + \cdots + T_0 + T_{-1} + \cdots + T_{-l} = \sum_{-l}^{l} T_i.$$

Diese Summe läßt sich aber durch folgende Betrachtung auf ein Integral, und zwar mit großer Annäherung, zurückführen.

Faßt man die Gleichung (6) als Gleichung einer Kurve auf, wobei T_x die Ordinate bedeutet, so kann die Fläche zwischen der Kurve, der Abszissenachse und den Ordinaten T_l und T_{-l} näherungsweise mittels der zu den Abszissen l, $l-1$, $l-2$, $\cdots -(l-1)$ gehörigen Ordinaten wie folgt ausgedrückt werden:

$$T_l + T_{l-1} + \cdots + T_{-(l-1)};$$

ein mit diesem gleichberechtigter Ausdruck lautet[1]):

1) Der erste Ausdruck entspricht der Summe der Rechtecke, welche erhalten werden, wenn man durch den Endpunkt jeder Ordinate bis an die benachbarte eine Parallele zu XX' nach rechts zieht; der zweite Ausdruck gibt die Summe der Rechtecke, welche durch das Ziehen der Parallelen nach links erhalten werden.

$$T_{l-1} + T_{l-2} + \cdots + T_{-l};$$

größere Genauigkeit wird erzielt, wenn man das Mittel beider Summen nimmt; demnach kann

$$\int_{-l}^{l} T_x dx = 2\int_0^l T_x dx = T_{l-1} + T_{l-2} + \cdots + T_{-(l-1)} + \frac{T_l + T_{-l}}{2},$$

also:

$$2\int_0^l T_x dx = \sum_{-l}^{l} T_i - \frac{T_l + T_{-l}}{2}$$

gesetzt werden, woraus[1])

---

1) Diese Formel, welche wir hier aus einem Falle der einfachen mechanischen Quadratur gewonnen haben, ist zuerst von Laplace (l. c.) mittels der Eulerschen Summenformel abgeleitet worden. Diese löst in aller Strenge das analytische Problem, die Summe $\sum_0^{l-1} T_i$ durch ein Integral auszudrücken, und lautet:

$$\sum_0^{l-1} T_i = \int_0^l T_x dx - \left\{ \frac{1}{2} T - \frac{B_1}{2!} T' + \frac{B_2}{4!} T''' - \frac{B_3}{6!} T^{\mathrm{v}} + \cdots \right\}_0^l;$$

darin bedeuten T', T''', $\cdots$ Ableitungen von T nach x und B_1, B_2, B_3, $\cdots$ die Bernoullischen Zahlen $\frac{1}{6}$, $\frac{1}{30}$, $\frac{1}{42}$, $\cdots$. Beschränkt man sich mit Rücksicht auf die Natur der hier vorliegenden Funktion auf das Glied $\frac{1}{2} T$, so ist

$$\sum_0^{l-1} T_i = \int_0^l T_x dx - \frac{1}{2} T_l + \frac{1}{2} T_0$$

$$\sum_0^{-(l-1)} T_i = \int_0^{-l} T_x dx - \frac{1}{2} T_{-l} + \frac{1}{2} T_0,$$

woraus

$$\sum_0^{l-1} T_i + \sum_0^{-(l-1)} T_i = 2\int_0^l T_x dx - \frac{T_l + T_{-l}}{2} + T_0;$$

nun ist aber

$$\sum_0^{l-1} T_i + \sum_0^{-(l-1)} T_i = \sum_{-l}^{l} T_i - (T_l + T_{-l}) + T_0,$$

daher

$$\sum_{-l}^{l} T_i = 2\int_0^l T_x dx + \frac{T_l + T_{-l}}{2}$$

wie oben.

$$\sum_{-l}^{l} T_l = 2\int_0^l T_x\,dx + \frac{T_l + T_{-l}}{2}$$

folgt.

Setzt man für T_x, T_l, T_{-l} die Werte der Gleichung (6) gemäß ein, so ergibt sich für P die Darstellung:

$$P = \frac{2}{\sqrt{2\pi spq}}\int_0^l e^{-\frac{x^2}{2spq}}dx + \frac{e^{-\frac{l^2}{2spq}}}{\sqrt{2\pi spq}},$$

welche noch dadurch vereinfacht werden kann, daß man

für $\frac{x}{\sqrt{2spq}}$ die neue Variable t,

für $\frac{l}{\sqrt{2spq}}$ den Buchstaben γ

einführt; alsdann wird

$$P = \frac{2}{\sqrt{\pi}}\int_0^\gamma e^{-t^2}\,dt + \frac{e^{-\gamma^2}}{\sqrt{2\pi spq}}. \tag{7}$$

77. Formulierung des Bernoullischen Theorems. Der Inhalt dessen, was in der gegenwärtigen Literatur unter dem Namen des Bernoullischen Theorems geführt wird, hat erst durch Laplace[1]) die endgültige Formulierung erhalten, in der wir das Theorem nunmehr vorführen. Man kann es in dem folgenden Satze zusammenfassen:

„Wenn über zwei entgegengesetzte Ereignisse E und $\bar{E}$, deren Wahrscheinlichkeiten p und q konstant sind, s Versuche angestellt werden, so ist unter allen möglichen Ergebnissen dasjenige am wahrscheinlichsten, in welchem das Verhältnis $m:n$ der Wiederholungszahlen von E und $\bar{E}$ dem Verhältnis $p:q$ der Wahrscheinlichkeiten gleich oder am nächsten ist.

Unter der Voraussetzung, daß die Zahl s der Versuche groß sei, drückt sich die Wahrscheinlichkeit dafür, daß die Wiederholungszahl m von E zwischen die Grenzen

$$sp - \gamma\sqrt{2spq} \quad \text{und} \quad sp + \gamma\sqrt{2spq}, \tag{8}$$

das Verhältnis $\frac{m}{s}$ dieser Wiederholungszahl zur Anzahl der Versuche oder die relative Häufigkeit des Eintreffens von E also zwischen die Grenzen

1) Théorie analyt., chap. III.

$$p - \gamma\sqrt{\frac{2pq}{s}} \quad \text{und} \quad p + \gamma\sqrt{\frac{2pq}{s}} \tag{9}$$

falle, durch

$$P = \frac{2}{\sqrt{\pi}}\int_0^{\gamma} e^{-t^2}\,dt + \frac{e^{-\gamma^2}}{\sqrt{2\pi spq}} \tag{7}$$

aus."

Um in den Sinn dieses weittragenden Satzes tiefer einzudringen, sei zunächst vorgreifend bemerkt, daß der Wert des ersten Gliedes auf der rechten Seite von (7) selbst für mäßige Werte von γ der Einheit sehr nahe ist und schon bei $\gamma = 4$ sich außerordentlich wenig von ihr unterscheidet; das zweite Glied liefert dann nur mehr einen sehr geringen Beitrag zu P, der unbeschadet der Richtigkeit der nachfolgenden Schlüsse vernachlässigt werden kann.

Aus den Ansätzen (8) und (9) ergibt sich nun:

„Die einer gegebenen Wahrscheinlichkeit P entsprechenden Grenzen (8) *der Wiederholungszahl m erweitern sich mit wachsender Anzahl der Versuche, jedoch nur im Verhältnis von $\sqrt{s}$."*

„Die einer gegebenen Wahrscheinlichkeit P entsprechenden Grenzen (9) *des Verhältnisses $\frac{m}{s}$, der relativen Häufigkeit des Auftretens von E, verengen sich mit wachsender Anzahl der Versuche und können durch entsprechende Vergrößerung von s beliebig eng gemacht werden."*

„Es ist möglich, die Zahl s der Versuche so groß zu wählen, daß man mit einer der Einheit beliebig nahen Wahrscheinlichkeit erwarten darf, es werde die relative Häufigkeit $\frac{m}{s}$ nicht mehr als innerhalb beliebig eng festgesetzter Grenzen von p abweichen."

Es muß ausdrücklich hervorgehoben werden, daß das Bernoullische Theorem nur Wahrscheinlichkeitsaussagen enthält und daher nur die Erwartungsbildung regelt, daß es also, wie jeder Satz der Wahrscheinlichkeitstheorie, über das wirkliche Geschehen keinen Aufschluß gibt. Wenn also mit einer der Einheit noch so nahen Wahrscheinlichkeit zu erwarten ist, $\frac{m}{s}$ werde innerhalb bestimmter enger, p umgebender Grenzen fallen, so kann eine wirklich ausgeführte Versuchsreihe doch ein Verhältnis $\frac{m}{s}$ ergeben, das weit über jene Grenzen hinausfällt, ohne daß hieraus ein Widerspruch mit unserem Theorem gefolgert werden dürfte.

78. Ausführung der Rechnungen. Die Berechnung von P bei gegebenem γ erfordert die Auswertung des Integrals

$$\Phi(\gamma) = \frac{2}{\sqrt{\pi}}\int_0^{\gamma} e^{-t^2}\,dt, \tag{10}$$

das eine höhere Transzendente vorstellt, die nur mittels unendlicher Rechenprozesse darstellbar ist. Auf die verschiedenen Methoden ihrer Berechnung soll hier nicht eingegangen werden.[1]) Wegen ihrer hohen praktischen Bedeutung sind Tafeln dieser Funktion berechnet worden; eine solche, Tafel I, ist am Schlusse des Buches mitgeteilt.[2])

Der zweite Teil von P hat bei einigermaßen hohem γ und großem s einen untergeordneten Einfluß und wird in manchen der neueren Schriften nicht geführt, so daß P durch das Integralglied allein ausgedrückt erscheint.[2])

Man könnte indessen, um bloß mit der Tafel auszukommen und doch den Wert dieses Gliedes zu berücksichtigen, dies durch eine Abänderung der Integralgrenze zu bewerkstelligen suchen. Es würde sich dann darum handeln, an die Stelle der Darstellung

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2spq}}} e^{-t^2} dt + \frac{e^{-\frac{l^2}{2spq}}}{\sqrt{2\pi spq}}$$

eine andere

$$P' = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l+\theta}{\sqrt{2spq}}} e^{-t^2} dt$$

mit möglichster Wahrung des Wertes zu setzen. Entwickelt man den letzten Ausdruck nach Potenzen von θ und bleibt bei der ersten Potenz stehen, so wird

$$P' = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2spq}}} e^{-t^2} dt + 2\theta \frac{e^{-\frac{l^2}{2spq}}}{\sqrt{2\pi spq}} + \cdots,$$

und soll $P' = P$ werden, so ist $\theta = \frac{1}{2}$ zu nehmen.[3]) Setzt man wieder

---

1) Näheres hierüber findet man in des Verf.s „Theorie der Beobachtungsfehler" (1891), p. 115—121, ferner bei H. Opitz, Die Kramp-Laplacesche Transzendente, Osterprogr. 1900 des Königstädt. Realgymn. zu Berlin.

2) Die Tafel ist durch sorgfältige Vergleichung der Tafeln von Encke (Berl. Astron. Jahrb. 1834), Bertrand (Calc. d. probab.), Glaisher (Philos. Mag. 42) und De Morgan (Encycl. Metrop. II, 1845) festgestellt worden. — Tafeln mit kleinerem Intervall von γ findet man bei Kämpfe in Wundts Philosoph. Studien 9 (1893); eine solche Tafel ist im Anschlusse an Tafel I am Ende des Buches mitgeteilt.

2) J. Bertrand, Calc. d. probab. (1889), p. 78 und 83; H. Poincaré, Calc. d. probab. (1896), p. 70.

3) Vgl. J. Eggenberger, Beiträge zur Darstellung des Bernoullischen Theorems, Berner Mitt. 50 (1894) und Zeitschr. f. Math. u. Ph. 45 (1900), p. 43.

$\frac{l+\frac{1}{2}}{\sqrt{2spq}} = \gamma$, so ändert sich der Ausdruck des Theorems dahin, daß

$$P = \frac{2}{\sqrt{\pi}}\int_0^{\gamma} e^{-t^2}\,dt$$

die Wahrscheinlichkeit für die Einhaltung der Grenzen

$$sp \mp \left(\gamma\sqrt{2spq} - \frac{1}{2}\right)$$

seitens m und der Grenzen

$$p \mp \left(\gamma\sqrt{\frac{2pq}{s}} - \frac{1}{2s}\right)$$

seitens $\frac{m}{s}$ bedeutet.

Handelt es sich darum, zu gegebenem P das zugehörige γ zu berechnen, so wird man aus Tafel I durch Interpolation die Wurzel der Gleichung

$$\Phi(\gamma) = P$$

bestimmen; es bleibt dann die an dieser Wurzel anzubringende Korrektur zu ermitteln, damit auch das zweite Glied von P Berücksichtigung finde. In den meisten Fällen wird man sich mit der Wurzel obiger Gleichung begnügen können. Der ganze Vorgang wird in der folgenden Nummer an einem besonderen Falle zur Erläuterung kommen.

79. Wahrscheinliche Abweichung. Diejenigen Grenzen von m, beziehungsweise von $\frac{m}{s}$, welchen die Wahrscheinlichkeit $P = \frac{1}{2}$ zukommt, so daß es ebenso wahrscheinlich ist, der zur Beobachtung kommende Wert werde zwischen sie fallen als er werde sie überschreiten, bezeichnet man als *wahrscheinliche Grenzen.* Die Hälfte ihres Intervalls heißt auch *wahrscheinliche Abweichung.*

Um sie zu finden, hat man jenen Wert von γ zu bestimmen, für welchen $P = \frac{1}{2}$ ist; er heiße ϱ, so daß

$$\frac{2}{\sqrt{\pi}}\int_0^{\varrho} e^{-t^2}\,dt + \frac{e^{-\varrho^2}}{\sqrt{2\pi spq}} = \frac{1}{2}$$

wird. Dann ist

$\varrho\sqrt{2spq}$ die wahrscheinliche Abweichung von m (und n),

$\varrho\sqrt{\frac{2pq}{s}}$ „ „ „ „ $\frac{m}{s}$ („ $\frac{n}{s}$).

Aus dem Ansatz

$$\frac{2}{\sqrt{\pi}}\int_0^{\varrho_0} e^{-t^2}\,dt = \frac{1}{2}$$

ergibt sich[1])

$$\varrho_0 = 0,4769362762;$$

setzt man sodann

$$\varrho = \varrho_0 + \delta,$$

so hat zur Bestimmung von δ die Gleichung zu dienen:

$$\frac{2}{\sqrt{\pi}}\int_0^{\varrho_0+\delta} e^{-t^2}\,dt + \frac{e^{-(\varrho_0+\delta)^2}}{\sqrt{2\pi spq}} = \frac{1}{2};$$

entwickelt man das erste Glied der linken Seite nach Potenzen von δ und bleibt bei der ersten Potenz stehen, so wird

$$\frac{2}{\sqrt{\pi}}\int_0^{\varrho_0} e^{-t^2}\,dt + \frac{2}{\sqrt{\pi}}e^{-\varrho_0^2}\cdot\delta + \frac{e^{-(\varrho_0+\delta)^2}}{\sqrt{2\pi spq}} = \frac{1}{2};$$

daraus aber ergibt sich mit Rücksicht auf die Bedeutung von ϱ_0 und darauf, daß $e^{-\varrho_0^2}$ nur sehr wenig verschieden ist von $e^{-(\varrho_0+\delta)^2}$,

$$\delta = -\frac{1}{2\sqrt{2spq}}.$$

Hiernach wäre

$$\varrho = 0,476936\cdots - \frac{1}{2\sqrt{2spq}} \tag{11}$$

zu setzen; die wahrscheinlichsten Grenzen von m würden

$$sp \mp \left(0,476936\sqrt{2spq} - \frac{1}{2}\right), \tag{12}$$

jene von $\frac{m}{s}$

$$p \mp \left(0,476936\sqrt{\frac{2pq}{s}} - \frac{1}{2s}\right) \tag{13}$$

sein. Indessen pflegt man den zweiten Teil von ϱ gegenüber dem

1) Nach der von H. Opitz (l. c.) ausgeführten Neuberechnung, welche in den letzten fünf Stellen von älteren Angaben differiert. — Wegen des häufigen Gebrauches setzen wir auch den Logarithmus von ϱ_0 her; er ist

$$\log \varrho_0 = \bar{9},6784587875.$$

ersten zu vernachlässigen und ϱ_0 für ϱ zu nehmen; dann entfällt in den beiden letzten Ansätzen das Glied $\frac{1}{2}$, beziehungsweise $\frac{1}{2s}$.

Um den ziffermäßigen Unterschied dieser zwei Rechnungsweisen zu beurteilen, möge die Frage nach der wahrscheinlichen Abweichung der Wiederholungszahl m von Wappen in $s = 4040$ Würfen mit einer Münze gegenüber der wahrscheinlichsten Zahl 2020, und nach der wahrscheinlichen Abweichung des Verhältnisses $\frac{m}{s}$ gegenüber dem normalen Werte $\frac{1}{2}$ auf beide Arten beantwortet werden.

Die strenge Rechnung gibt für die erstgedachte Abweichung

$$0{,}4769\sqrt{2020} - \frac{1}{2} = 20{,}9\cdots,$$

also (bei einer geringfügigen Erhöhung von P über $\frac{1}{2}$) 21, für die Abweichung $\frac{m}{s} - \frac{1}{2}$

$$\frac{0{,}4769}{\sqrt{8080}} - \frac{1}{8080} = 0{,}052935;$$

nach der abgekürzten Rechnung erhält man

21,4, also (bei einer geringfügigen Verminderung von P) 21, beziehungsweise 0,053058.

80. Mittlere Abweichung. Unter der mittleren Abweichung der Wiederholungszahl m, respektive der relativen Häufigkeit $\frac{m}{s}$ von dem zugehörigen wahrscheinlichsten Werte sp, beziehungweise p, versteht man die Quadratwurzel aus dem mittleren Quadrat dieser Abweichung.

Nach einem in Nr. 55 entwickelten Satze ist der Mittelwert einer vom Zufall abhängigen Größe gleich der Summe der Produkte ihrer Einzelwerte mit den zugehörigen Wahrscheinlichkeiten. Setzt man also $m - sp = l$, so ist

$$\mathfrak{D}(l^2) = \sum_{m=s}^{0} (m - sp)^2 \frac{s!}{m!\,n!} p^m q^n;$$

denn $\frac{s!}{m!\,n!} p^m q^n$ ist die Wahrscheinlichkeit, daß sich E m-mal wiederholen, also die Abweichung $m - sp$ einstellen werde.

Die weitere Entwicklung des obigen Ausdrucks gibt

$$\mathfrak{D}(l^2) = \sum_{s}^{0} \frac{s!}{m!\,n!} m^2 p^m q^n - 2sp \sum_{s}^{0} \frac{s!}{m!\,n!} m p^m q^n + s^2 p^2 \sum_{s}^{0} \frac{s!}{m!\,n!} p^m q^n;$$

nun ist

$$\sum_s^0 \frac{s!}{m!\,n!} p^m q^n = (p+q)^s = 1;$$

verallgemeinert man diesen Ansatz mittels der Variablen t zu

$$\sum_s^0 \frac{s!}{m!\,n!} (pt)^m q^n = (pt+q)^s$$

und differenziert ihn dann zweimal nach t, so ergeben sich die Gleichungen:

$$\sum_s^0 \frac{s!}{m!\,n!} m p (pt)^{m-1} q^n = sp\,(pt+q)^{s-1}$$

$$\sum_s^0 \frac{s!}{m!\,n!} m(m-1) p^2 (pt)^{m-2} q^n = sp^2 (s-1)(pt+q)^{s-2};$$

die Substitution $t = 1$ liefert:

$$\sum_s^0 \frac{s!}{m!\,n!} m p^m q^n = sp,$$

$$\sum_s^0 \frac{s!}{m!\,n!} m(m-1) p^m q^n$$

$$= \sum_s^0 \frac{s!}{m!\,n!} m^2 p^m q^n - \sum_s^0 \frac{s!}{m!\,n!} m p^m q^n = sp^2(s-1),$$

woraus

$$\sum_s^0 \frac{s!}{m!\,n!} m^2 p^m q^n = sp^2(s-1) + sp$$

folgt. Mit diesen Werten der drei Summen erhält man

$$\mathfrak{D}(l^2) = spq.$$

Hiernach ist die mittlere Abweichung der Wiederholungszahl m von sp gleich

$$\sqrt{spq} \tag{14}$$

und, wie eine einfache Überlegung zeigt,

$$\sqrt{\frac{pq}{s}} \tag{15}$$

die mittlere Abweichung des Verhältnisses $\frac{m}{s}$ von p.

Würde man das approximative Gesetz der Wahrscheinlichkeit der Abweichungen, wie es durch die Gleichung (6), Nr. 75 dargestellt ist, zur Berechnung von $\mathfrak{D}(l^2)$ heranziehen, hierbei l als eine stetige Größe behandeln und die Integration mit Rücksicht auf die außerordentlich rasche Abnahme der Exponentialgröße auf das Intervall $(-\infty, \infty)$ ausdehnen, so ergäbe sich:

$$\mathfrak{D}(l^2) = \frac{1}{\sqrt{2\pi spq}} \int_{-\infty}^{\infty} e^{-\frac{l^2}{2spq}} l^2 \, dl;$$

setzt man hierin $\frac{l}{\sqrt{2spq}} = t$, so wird weiter

$$\mathfrak{D}(l^2) = \frac{2spq}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt = spq,$$

weil der Wert des letzten Integrals (s. Nr. 12) $\frac{\sqrt{\pi}}{2}$ ist. Es ergibt sich also auf diesem Wege dasselbe Resultat, zu dem die strenge Rechnung geführt hat. Man kann in diesem Umstande eine Bestätigung dafür erblicken, daß das approximative Gesetz wohl geeignet ist, die Wahrscheinlichkeit der Abweichungen auf deren ganzem Gebiete darzustellen.

81. Mittelwert der Abweichung und durchschnittliche Abweichung. Unter der ersteren Benennung hat man sinngemäß $\mathfrak{D}(l)$ zu verstehen, wobei $l = m - sp$ ist. Als durchschnittliche Abweichung hingegen bezeichnet man $\mathfrak{D}(|l|)$, also den Mittelwert der absoluten Werte von $m - sp$.

Für $\mathfrak{D}(l)$ ergibt sich durch eine einfache Rechnung der Wert Null; denn es ist

$$\mathfrak{D}(l) = \sum_{s}^{0} \frac{s!}{m!\,n!}(m - sp)p^m q^n = \sum_{s}^{0} \frac{s!}{m!\,n!} m p^m q^n - sp \sum_{s}^{0} \frac{s!}{m!\,n!} p^m q^n,$$

also nach Nr. 80

$$\mathfrak{D}(l) = sp - sp = 0. \tag{16}$$

Für $\mathfrak{D}(|l|)$ hat man zunächst den folgenden Ansatz:

$$\mathfrak{D}(|l|) = \sum_{m=s}^{sp} \frac{s!}{m!\,n!}(m - sp)p^m q^n + \sum_{sp}^{0} \frac{s!}{m!\,n!}(sp - m)p^m q^n;$$

ersetzt man in der ersten Summe $m - sp$ durch das gleichwertige $mq - np$, so löst sie sich auf in

$$q\sum_{s}^{sp}\frac{s!}{m!\,n!}\,m\,p^m q^n - p\sum_{s}^{sp}\frac{s!}{m!\,n!}\,n p^m q^n;$$

es ist aber

$$\frac{\partial}{\partial p}\sum_{s}^{sp}\frac{s!}{m!\,n!}p^m q^n = \sum_{s}^{sp}\frac{s!}{m!\,n!}\,m\,p^{m-1}q^n$$

$$\frac{\partial}{\partial q}\sum_{s}^{sp}\frac{s!}{m!\,n!}p^m q^n = \sum_{s}^{sp}\frac{s!}{m!\,n!}\,n\,p^m q^{n-1},$$

daher

$$\sum_{s}^{sp}\frac{s!}{m!\,n!}(m-sp)p^m q^n$$

$$= pq\left\{\frac{\partial}{\partial p}\sum_{s}^{sp}\frac{s!}{m!\,n!}p^m q^n - \frac{\partial}{\partial q}\sum_{s}^{sp}\frac{s!}{m!\,n!}p^m q^n\right\}.$$

Man hat also

$$p^s + \frac{s!}{(s-1)!\,1!}p^{s-1}q + \frac{s!}{(s-2)!\,2!}p^{s-2}q^2 + \cdots + \frac{s!}{sp!\,sq!}p^{sp}q^{sq}$$

zuerst nach p, dann nach q zu differenzieren und das zweite Resultat von dem ersten in Abzug zu bringen; das Ergebnis dieser Rechnung ist

$$s\,\frac{s!}{sp!\,sq!}p^{sp}q^{sq} = s\,T_0,$$

so daß der erste Teil von $\mathfrak{D}(|l|)$ den Wert $spqT_0$ hat. Eine einfache Überlegung lehrt, daß dem zweiten Teile derselbe Wert zukommt, daß also

$$\mathfrak{D}(|l|) = 2\,spq\,T_0.$$

Ersetzt man T_0 durch den in Nr. 74 gefundenen Näherungswert, so ergibt sich[1])

$$\mathfrak{D}(|l|) = \sqrt{\frac{2\,spq}{\pi}}. \tag{17}$$

Selbstverständlich ist

$$\mathfrak{D}\left(\left|\frac{m}{s}-p\right|\right) = \sqrt{\frac{2\,pq}{\pi s}}. \tag{18}$$

Dieselben Werte von $\mathfrak{D}(l)$ und $\mathfrak{D}(|l|)$ erhält man, wenn man den am Schlusse der vorigen Nummer bezüglich $\mathfrak{D}(l^2)$ eingeschlagenen Weg befolgt.

1) Vgl. hiermit die Ableitung J. Bertrands, Calc. d. probab., p. 82.

82. Beispiel XLIV. *Mit einer Münze werden 4040 Würfe ausgeführt. Wie groß ist die Wahrscheinlichkeit, daß die Wiederholungszahl von Wappen nicht um mehr als 28 von der wahrscheinlichsten Zahl 2020 abweichen, also zwischen 1992 und 2048 liegen werde?*

Gemäß den Formeln (8) und (7) in Nr. 77 ist

$$\gamma = \frac{28}{\sqrt{2020}} = 0{,}623,$$

$$P = \Phi(0{,}623) + \frac{e^{-0{,}623^2}}{\sqrt{2020\pi}};$$

der Tafel I entnimmt man

$$\Phi(0{,}62) = 0{,}6194114,$$

daraus ergibt sich durch Benützung der ersten und zweiten Differenz

$$\Phi(0{,}623) = 0{,}6217119;$$

der zweite Teil liefert in logarithmischer Ausrechnung den Beitrag

$$0{,}0085151;$$

mithin ist

$$P = 0{,}6302270.$$

Bei alleiniger Benützung der Tafel I hätte man nach dem in Nr. 78 entwickelten Verfahren mit

$$\gamma = \frac{28{,}5}{\sqrt{2020}} = 0{,}63412$$

zu rechnen und fände durch Interpolation

$$P = \Phi(0{,}63412) = 0{,}6301641,$$

also um 0,0000629 weniger als durch strenge Rechnung.

83. Beispiel XLV. *Welches sind die wahrscheinlichen Grenzen für die Wiederholungszahl einer bezeichneten Nummer in 2854 Lotterieziehungen?*

Da die Wahrscheinlichkeit für das Erscheinen einer bestimmten Nummer in einer Ziehung $p = \frac{1}{18}$ ist, so hat man nach den Formeln von Nr. 79 für die wahrscheinliche Abweichung der Wiederholungszahl den Wert

$$0{,}476936\sqrt{2 \cdot 2854 \cdot \frac{1}{18} \cdot \frac{17}{18}} = 8{,}2 \cdots;$$

da die wahrscheinlichste Wiederholungszahl $\frac{1}{18} \cdot 2854$, d. i. 158 beträgt,

so besteht die Wahrscheinlichkeit $\frac{1}{2}$ dafür, es werde sich die bezeichnete Nummer nicht weniger als 150 mal und nicht öfter als 166 mal wiederholen.[1])

84. Beispiel XLVI. *Wie viele Versuche sind erforderlich, um mit der Wahrscheinlichkeit P erwarten zu dürfen, daß die Wiederholungszahl eines Ereignisses, dem im einzelnen Versuche die Wahrscheinlichkeit p zukommt, nicht mehr als um $\mathfrak{p}$ Prozent von dem wahrscheinlichsten Werte abweichen werde?*

Die Lösung dieser Aufgabe vollzieht sich mittels der beiden Gleichungen

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2}\,dt,$$

$$\gamma\sqrt{2spq} = \frac{\mathfrak{p}}{100}\,s;$$

nachdem man aus der ersten γ bestimmt hat, ergibt sich aus der zweiten:

$$s = 2pq\left(\frac{100\,\gamma}{\mathfrak{p}}\right)^2.$$

Will man z. B. mit der Wahrscheinlichkeit $\frac{999}{1000}$ erwarten dürfen, daß die Zahl der weißen Kugeln, die man aus einer Urne mit fünf weißen und einer schwarzen Kugel in s Ziehungen erzielt, höchstens um 5% von ihrem wahrscheinlichsten Werte $\frac{s}{6}$ abweiche, so suche man zuerst mit Hilfe der Taf. I den Wert γ, der

$$\frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2}\,dt = 0{,}999$$

ergibt; durch Benützung der ersten Differenz findet man

$$\gamma = 2{,}32 + \frac{0{,}0000345}{0{,}0000507}\,0{,}01 = 2{,}32684;$$

mit diesem Werte berechnet sich dann

$$s = 2 \cdot \frac{5}{6} \cdot \frac{1}{6}\left(\frac{232{,}684}{5}\right)^2 = 601{,}5\cdots;$$

hiernach sind 602 Ziehungen für den gedachten Zweck erforderlich.

---

1) Wegen der Abrundung des Resultates auf eine ganze Zahl ist die Wahrscheinlichkeit um ein Geringes von $\frac{1}{2}$ verschieden, hier etwas kleiner.

85. Begriff der Präzision. Zwei Versuchsreihen, mögen sie sich auf dieselbe oder auf verschiedene Materien beziehen, sind in einem Punkte miteinander vergleichbar, nämlich in bezug auf die Weite der Grenzen, innerhalb welcher sie die Abweichung $\frac{m}{s} - p$ mit *gegebener* Wahrscheinlichkeit P erwarten lassen.

Um diese Vorstellung zu präzisieren, denken wir uns zwei Beobachtungs- oder Versuchsreihen; die eine beziehe sich auf die Ereignisse E, $\bar{E}$, denen im ganzen Verlaufe der s Versuche die Wahrscheinlichkeiten p, q zukommen; die andere betreffe die Ereignisse E' $\bar{E}'$ mit den Wahrscheinlichkeiten p', q' und umfasse s' Versuche; die Wiederholungszahl von E sei m, die von E' sei m'. Alsdann ist mit der Wahrscheinlichkeit[1])

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt$$

zu erwarten, daß einerseits

$$-\gamma \sqrt{\frac{2pq}{s}} < \frac{m}{s} - p < \gamma \sqrt{\frac{2pq}{s}}$$

und daß andererseits

$$-\gamma \sqrt{\frac{2p'q'}{s'}} < \frac{m'}{s'} - p' < \gamma \sqrt{\frac{2p'q'}{s'}};$$

die Weite dieser Intervalle ist

$$2\gamma \sqrt{\frac{2pq}{s}}, \quad 2\gamma \sqrt{\frac{2p'q'}{s'}}$$

und hängt bei gegebenem P, also bei festem γ, nur von dem Wurzelausdruck ab. Je enger nun das Intervall, je näherliegend an p man das Verhältnis $\frac{m}{s}$ mit gegebener Wahrscheinlichkeit zu erwarten hat, für desto genauer wird man die betreffenden Versuche erklären. Die Genauigkeit wächst also mit der Abnahme von $\sqrt{\frac{2pq}{s}}$, oder sie wächst mit dem Wachsen von

$$h = \sqrt{\frac{s}{2pq}},$$

und man kann übereinkommen, diesen Ausdruck geradezu als das Maß der *Präzision* der betreffenden Versuchsreihe aufzufassen.

1) Behielte man den vollständigen Ausdruck (7), Nr. 76, für P bei, so würde dies auf die nachfolgende Betrachtung einen ziffermäßig nur geringfügigen Einfluß üben.

Die Präzision wächst sonach wie die Quadratwurzel aus der Anzahl der Versuche und ist um so größer, je kleiner das Produkt pq, je mehr p und q von $\frac{1}{2}$ abweichen. In dem erklärten Sinne haben also Versuche, welche sich mit dem Aufwerfen einer Münze vergleichen lassen, die kleinste Präzision, $h = \sqrt{2s} = 1{,}414 \cdot \sqrt{s}$; dem gegenüber ist beispielsweise die Präzision für das Erscheinen einer bestimmten Nummer in s Ziehungen der gewöhnlichen Lotterie:

$$h' = \sqrt{\frac{s}{2 \cdot \frac{17}{18^2}}} = 2{,}850 \sqrt{s},$$

also zweimal so groß.

Im Laufe der vorangegangenen Untersuchungen sind für die wahrscheinliche, für die mittlere und für die durchschnittliche Abweichung der relativen Häufigkeit $\frac{m}{s}$ von der Wahrscheinlichkeit p die Ausdrücke

$$\varrho\sqrt{\frac{2pq}{s}}, \quad \sqrt{\frac{pq}{s}}, \quad \sqrt{\frac{2pq}{\pi s}}$$

[s. die Ansätze (13), (15), (18)] gefunden worden; dieselben sind durchweg der Präzision h umgekehrt proportional; demnach kann die Präzision auch nach diesen Größen beurteilt werden.

86. Wiederholte Versuchsreihen. Solange nur *eine* Versuchsreihe vorliegt, läßt sich über ihr Ergebnis vom Standpunkte der Wahrscheinlichkeitstheorie keine andere Aussage machen als die, daß es eine bestimmte oder eine innerhalb vorgegebener Grenzen liegende Abweichung vom wahrscheinlichsten Ergebnis mit dieser oder jener Wahrscheinlichkeit erwarten lasse.

Zu weitergehenden Schlüssen und Betrachtungen gibt der Fall Anlaß, daß über zwei Ergebnisse E, $\bar{E}$ mit festen Wahrscheinlichkeiten p, q mehrere, sagen wir z, Beobachtungsreihen gleichen Umfanges, also auch gleicher Präzision, angestellt worden sind. Jede Reihe bestehe aus s Versuchen, und es habe die erste m_1, die zweite $m_2, \cdots$ die letzte m_z als Wiederholungszahl von E ergeben. Dann kann von einer *wahrscheinlichsten* Gruppierung, Verteilung oder *Dispersion* der Quotienten

$$\frac{m_1}{s}, \quad \frac{m_2}{s}, \quad \cdots \quad \frac{m_z}{s}$$

um den Wert p gesprochen werden.

Da nach dem approximativen Wahrscheinlichkeitsgesetz

$$T = \frac{1}{\sqrt{2\pi spq}} e^{-\frac{t^2}{2spq}}$$

eine positive Abweichung dieselbe Wahrscheinlichkeit hat wie eine gleich große negative, so ist es am wahrscheinlichsten, daß sich ebenso viele der Werte $\frac{m_i}{s}$ unter p wie über p stellen werden.

Nach dem Bernoullischen Theorem besteht ferner eine bestimmte Wahrscheinlichkeit $P = \Phi(\gamma)$ dafür, daß die Abweichung $\frac{m_i}{s} - p$ dem Betrage nach nicht größer sei als $\gamma\sqrt{\frac{2pq}{s}}$; daraus folgt, daß zP die wahrscheinlichste Zahl derjenigen Quotienten $\frac{m_i}{s}$ ist, welche sich in das Intervall zwischen $p - \gamma\sqrt{\frac{2pq}{s}}$ und $p + \gamma\sqrt{\frac{2pq}{s}}$ einstellen.

Dazu ist aber zu bemerken, daß es, je größer z, um so weniger wahrscheinlich ist, es werde sich genau diese wahrscheinlichste Zahl zP ergeben; vielmehr muß eine Abweichung von dieser Zahl erwartet werden, und über diese Abweichung kann auf Grund des Bernoullischen Theorems wieder eine Wahrscheinlichkeitsaussage gemacht werden. Es ist nämlich mit der Wahrscheinlichkeit

$$\Pi = \frac{2}{\sqrt{\pi}} \int_0^\delta e^{-t^2} dt$$

zu erwarten, daß die wirklich zur Beobachtung kommende Anzahl der Quotienten $\frac{m_i}{s}$, welche zwischen die oben angeführten Grenzen fallen, liegen werde zwischen

$$zP - \delta\sqrt{2zP(1-P)} \quad \text{und} \quad zP + \delta\sqrt{2zP(1-P)}.$$

Zur Erläuterung diene folgendes Beispiel. Die wahrscheinlichste Wiederholungszahl der Nummer 1 in $s = 2854$ Lotterieziehungen ist $\frac{1}{18} \cdot 2858$, d. i. 158. Die gleiche Zahl gilt für jede der 89 übrigen Nummern, und man kann die Ziehungen wie $z = 90$ gleich umfangreiche Beobachtungsreihen auffassen, indem man sie einmal auf das Erscheinen der Nummer 1, ein zweites Mal auf das Erscheinen der Nummer 2 usw. hin prüft.

Die Wahrscheinlichkeit, daß die Wiederholungszahl einer bestimmten Nummer zwischen die Grenzen 156 und 160 (mit Einschluß derselben) falle, ist

$$P = \Phi(\gamma)$$

mit

$$\gamma = \frac{2}{\sqrt{2 \cdot 2854 \cdot \frac{1}{18} \cdot \frac{17}{18}}} = 0{,}11557,$$

also

$$P = \Phi(0{,}11557) = 0{,}1298;$$

daraus ergibt sich die wahrscheinlichste Anzahl derjenigen Nummern, welche diese Grenzen einhalten,

$$zP = 90 \cdot 0{,}1298 = 11, \cdots;$$

die Wahrscheinlichkeit, daß von den 90 Nummern nicht weniger als 10 und nicht mehr als 12 die erwähnten Grenzen einhalten werden, ist

$$\Pi = \Phi(\delta)$$

mit

$$\delta = \frac{1}{\sqrt{2 \cdot 90 \cdot 0{,}1298 \cdot 0{,}8702}} = 0{,}219,$$

also

$$\Pi = 0{,}24321.$$

87. Beziehung zwischen Wahrscheinlichkeitsrechnung und Wirklichkeit. Gesetz der großen Zahlen. Wir treten nun an eine Frage heran, von deren Beantwortung es abhängen wird, ob die Wahrscheinlichkeitsrechnung lediglich zu den rein spekulativen Zweigen der Mathematik gehört oder ob sich angewandte Gebiete auf sie gründen lassen. Es handelt sich darum, ob den Sätzen der Wahrscheinlichkeitsrechnung irgendein Korrelat im wirklichen Geschehen entspricht.

Wohl stützt sich die Bildung des Wahrscheinlichkeitsbruches auf die Kenntnis gewisser Umstände, von denen anzunehmen ist, daß sie auf das Geschehen einen Einfluß üben, und seine Struktur hängt mit diesem Wissen derart zusammen, daß man ihn als Maß für die Erwartung eines bestimmten Geschehens betrachtet. Ob aber das wirkliche Geschehen zu den Ergebnissen der logischen Deduktion, als welche sich die Wahrscheinlichkeitsaussagen darstellen, in einer Beziehung steht, kann nur durch Befragen der Wirklichkeit selbst entschieden werden.

Ein einzelner Versuch, eine einmalige Realisierung der allgemeinen Bedingungen kann einen Aufschluß nicht gewähren; denn eine der Möglichkeiten *muß* sich einstellen, und darüber hinaus läßt sich keine weitere Aussage machen. Kommt dem einen möglichen Ereignis E eine der Einheit, dem andern $\bar{E}$ demgemäß eine der Null noch so

nahe Wahrscheinlichkeit zu, so kann der Erfolg trotz der Stärke der auf E gerichteten Erwartung doch $\bar{E}$ bringen; und sowie das etwaige Eintreffen von E nicht als Beweis *für* eine reale Bedeutung des Wahrscheinlichkeitsbruches hingestellt werden kann, wird das Eintreffen von $\bar{E}$ nicht als Beweis *gegen* eine solche geführt werden dürfen.

Anders gestaltet sich die Sachlage, wenn über eine unveränderliche Materie eine große Zahl von Realisierungsergebnissen vorliegt; um an einen bestimmten Fall zu denken, wenn aus einer Urne, die mit weißen und schwarzen Kugeln in solchem Verhältnisse gefüllt ist, daß dem Eintreffen einer weißen (E) die Wahrscheinlichkeit p, dem einer schwarzen ($\bar{E}$) die Wahrscheinlichkeit $q = 1 - p$ zukommt, eine große Anzahl s von Ziehungen ausgeführt wurde und m-mal weiß und $n = s - m$-mal schwarz sich ergab.

In bezug auf einen solchen Tatbestand gibt die Wahrscheinlichkeitsrechnung in dem Bernoullischen Theorem eine Aussage, aus der sich ein für die ganze Frage maßgebender Schluß ziehen läßt. Das Theorem bestimmt nämlich die Wahrscheinlichkeit P dafür, daß der Unterschied zwischen der relativen Häufigkeit $\frac{m}{s}$ des Auftretens von E und seiner Wahrscheinlichkeit p dem Betrage nach nicht größer ausfällt als $\gamma\sqrt{\frac{2pq}{s}}$; nun kann γ so eingerichtet werden, daß P der Einheit so nahe kommt, als man will, und es reicht dazu ein mäßiger Wert von γ hin; bezeichnet man also die relative Häufigkeit von E in s Versuchen mit $H(E, s)$, so führt das Bernoullische Theorem zu dem Schlusse: Man hat mit einer der Einheit beliebig nahen Wahrscheinlichkeit zu erwarten, daß

$$\lim_{s=\infty} \{H(E,s) - p\} = \lim \gamma \sqrt{\frac{2pq}{s}} = 0$$

werde.

Über den Verlauf dieses Grenzprozesses, für den doch nur Wahrscheinlichkeit besteht, wenn sie auch noch so groß ist, läßt sich aber eine wichtige Bemerkung machen. Der rein mathematische Grenzprozeß verläuft so, daß man mit Sicherheit angeben kann, von einem Werte s der Variablen angefangen werde der Unterschied zwischen der Funktion und ihrer Grenze unter dem und dem Betrage bleiben. Hier jedoch kann immer wieder das Gegenteil des erwarteten eintreffen, d. h. die relative Häufigkeit nach einer Versuchszahl s außerhalb der mit P erwarteten Grenzen fallen; dies hat zur Folge, daß einmal schon für ein endliches s die *Gleichung*

$$H(E,s) = p$$

sich einstellen, daß andererseits trotz erheblicher Ungleichheit im Sinne $s' < s''$ doch

$$| H(E,s'') - p | > | H(E,s') - p |$$

stattfinden kann.

In Worten ausgedrückt: Das Bernoullische Theorem läßt nur erwarten, daß mit wachsender Versuchszahl die absolute Differenz zwischen der relativen Häufigkeit und der Wahrscheinlichkeit von E *im allgemeinen* abnehmen werde.

Damit ist eine Handhabe zur Befragung der Wirklichkeit geboten: Zeigt der tatsächliche Verlauf des Geschehens diese Eigenschaft, so ist damit die Konkordanz zwischen den Sätzen der Wahrscheinlichkeitsrechnung und der Wirklichkeit nachgewiesen.

Es soll an dieser Stelle die entscheidende Tatsache festgestellt werden, daß *in all den Fällen, wo eine Prüfung der Wirklichkeit in dem oben erklärten Sinne erfolgt ist und wo die Voraussetzungen für eine begründete Aufstellung des Wahrscheinlichkeitsbruches vorhanden waren, eine so weitgehende Annäherung der relativen Häufigkeit an die Wahrscheinlichkeit festgestellt werden konnte, daß man sagen darf, der theoretische Satz*

$$\lim_{s=\infty} H(E,s) = p \tag{1}$$

finde in der Wirklichkeit in solchem Grade Bestätigung, als dies von einer Anwendung der Ergebnisse reinen Denkens auf die verwickelten Vorgänge des Geschehens erwartet werden kann.

Der im Wege der Erfahrung vom Boden der Theorie auf den Boden der Wirklichkeit übertragene Satz (1), wenn man hier seinen Sinn richtig deutet, macht den Inhalt dessen aus, was man gemeinhin als das *Gesetz der großen Zahlen* bezeichnet. Mit diesem Gesetze rückt das, was wir im einzelnen Falle zufällig nennen, in der großen Menge dem ursächlich Gesetzmäßigen so nahe, daß es für praktische Zwecke genau genug als vom Charakter des Zufälligen befreit erachtet werden kann. In der Tat arbeiten auf zufällige Massenerscheinungen gegründete Unternehmungen mit derselben Sicherheit wie andere geschäftliche Einrichtungen, bei denen es nicht üblich ist, vom Zufall zu reden.

Über die Stellung, welche dem Satze in jüngster Zeit im System der Wahrscheinlichkeitsrechnung gegeben wird, sei folgendes bemerkt:

Bruns[1]) stellt schon an den Eingang des Lehrgebäudes eine Aussage, die er als Satz von der Ausgleichung des Zufalls oder als *Satz von der gleichmäßigen Erschöpfung der möglichen Fälle* und als die-

1) Wahrscheinlichkeitsrechnung und Kollektivmaßlehre, Leipzig 1906, p. 13

jenige „Voraussetzung" bezeichnet, auf der alle ernsthaften Anwendungen der Wahrscheinlichkeitsrechnung beruhen. Die Namengebung stammt daher, daß Bruns die einzelnen möglichen Fälle in Betracht zieht. Führt eine Urteilsmaterie auf n gleichberechtigte Fälle A_1, $A_2, \cdots A_n$, so formuliert er seinen Satz dahin, daß

$$\lim_{s=\infty} H(A_i, s) = \frac{1}{n} \qquad (i = 1, 2, \cdots n). \tag{2}$$

Ebenso erscheint in Brendels Vorlesungen[1]) der Satz unter den Grundlagen der Wahrscheinlichkeitsrechnung und wird hier in einer nach meinem Dafürhalten zu sehr spezialisierten Form als „Axiom" hingestellt. Wenn dem Ereignis E unter m gleichberechtigten g Fälle günstig sind, und wenn in n Reihen von je m Versuchen das Ereignis $\gamma_1, \gamma_2, \cdots \gamma_n$-mal wirklich eingetreten ist, so gilt als Axiom:

$$\lim_{n=\infty} \frac{\sum_{i=1}^{n} \gamma_i}{n} = g. \tag{3}$$

Wenn nun auch keinerlei logische Schlüsse und keinerlei naturphilosophische Betrachtungen die Notwendigkeit des in Rede stehenden Satzes zu *beweisen* vermögen, so muß sich sein durch die Erfahrung erhärteter Bestand doch *erklären* lassen. Die Erklärung liegt, wenn man sich auf das Urnenschema bezieht, in der *Existenz* der möglichen Fälle und in einer so gearteten *Verwirklichung* der allgemeinen Bedingungen, daß sich der Bereich der Verwirklichungsvorgänge auf die Fälle nahezu gleichmäßig verteilt.[2])

Daß die bloße Existenz der Fälle zur Erklärung nicht ausreicht, lehrt die nachfolgende Überlegung. Wären die Kugeln in unserer Urne so angeordnet, daß die Hauptmenge der schwarzen zu unterst liegt und daß Ziehung und Mischung sich immer nur auf die obersten

---

1) Wahrscheinlichkeitsrechnung mit Einschluß der Anwendungen (Autogr.), Göttingen 1907, p. 18.

2) Eine befremdende Ausführung über den Zusammenhang zwischen Wahrscheinlichkeit und Wirklichkeit findet sich in der an historischen Daten reichen Abhandlung P. Mansions: „Sur la portée objective du calcul des probabilités", Bull. de l'Acad. de Belgique, 1903, p. 1256 sq. Nachdem die zutreffende Bemerkung gemacht worden, daß man in Wirklichkeit niemals eine objektive Wahrscheinlichkeit mit absoluter Genauigkeit anzugeben vermöge, heißt es weiter: „Noch mehr; so fremdartig es erscheinen mag, die Wahrscheinlichkeitsrechnung würde ihre Existenz einbüßen, sie hätte keine Daseinsberechtigung mehr, wenn die objektive Wahrscheinlichkeit eines Ereignisses einen völlig bestimmten Wert hätte. Welchen Sinn kann man in der Tat der Aussage geben: Die Wahrscheinlichkeit des Eintreffens von 5 beim Würfeln mit einem bestimmten Würfel sei beständig $\frac{1}{6}$? Doch keinen andern als: In 6 Würfen erscheint 5 einmal, in 12 zweimal, in 600 hundertmal usw. Nicht genug an dem, vom Beginne des Spieles

Schichten erstrecken, dann wird $H(E, s)$ bei zunehmendem s im allgemeinen beständig von p erheblich verschieden sein.

Ein Bild noch möge zur Beleuchtung des Vorgangs, den man als Ausgleichung des Zufalls bezeichnet, beitragen. Bringt man in ein Gefäß größere Mengen von Roggen- und Gerstenkörnern und rührt die ganze Masse durch, so wird, je länger die Prozedur dauert, eine um so gleichmäßigere Verteilung beider Sorten von Körnern sich einstellen und das Gemenge der Homogenität in dem Sinne zustreben, daß in Teilmengen, die an beliebigen Stellen herausgegriffen werden, die relative Häufigkeit beider Körnersorten nahezu konstant wird.

Wenn einmal eine Versuchsreihe von beträchtlichem Umfang eine erhebliche Abweichung von dem durch das Gesetz der großen Zahlen ausgedrückten Verhalten zeigen sollte, so geben die voranstehenden Betrachtungen Anhaltspunkte dafür, nach welcher Richtung Vermutungen über den Grund dieser Erscheinung aufgestellt werden könnten. Einmal kann er darin liegen, daß die vorausgesetzte Gleichmöglichkeit der Fälle in den objektiven Umständen nicht begründet ist; eine neue Versuchsreihe würde dann voraussichtlich zu einer ähnlichen Abweichung führen. Der Grund kann ferner darin liegen, daß die Durchführung der Versuche nicht der Forderung genügte, allen Fällen gewissermaßen gleiche Gelegenheit, in die Erscheinung zu treten, geboten zu haben; von einer Wiederholung der Versuchsreihe ist dann nicht anzunehmen, daß sie eine der früheren ähnliche Erscheinung darbieten werde. Schließlich kann es unter Umständen geboten sein, auch darnach zu forschen, ob die Sätze der Wahrscheinlichkeitsrechnung richtig angewendet worden sind.

Untersuchungen über die Geltung des Gesetzes der großen Zahlen sind schon vielfach angestellt und veröffentlicht worden; zum Teil stützten sie sich auf bereits vorhandene Erfahrungsreihen, zum Teil auf eigens zu diesem Zwecke ausgeführte Versuchsreihen. Wiederholt sind auch Materien der geometrischen Wahrscheinlichkeit zu solchen Versuchen herangezogen worden. Die Frage, ob dies auch zulässig sei, würde hier nicht berührt werden, wenn nicht Zweifel hierüber ausgesprochen worden wären. Cesàro[1]) berichtet über eine Meinung G. Vivantis, der behauptet hat, daß bei Problemen mit unendlich

---

an müßte sich 5 in jeder Serie von 6 Würfen immer an derselben Stelle ereignen; denn sonst gäbe es bei einer entsprechend ausgeführten Unterteilung der 600 Würfe sechsgliedrige Gruppen ohne 5. Wenn sich aber 5 mit dieser absoluten Regelmäßigkeit zutrüge, dann wäre nichts Wahrscheinliches und Unwahrscheinliches mehr in seinem Eintreffen, man wäre in Gewißheit darüber, daß es sich bei dem so und sovielten Wurfe zutrage; dies wäre ein Naturgesetz." Wie dies alles aus der Existenz einer bestimmten objektiven Wahrscheinlichkeit gefolgert werden soll, ist nicht zu erkennen.

1) l. c., p. 49—62.

vielen Möglichkeiten die Brücke zwischen Theorie und Wirklichkeit fehle, die sich für die andern Probleme im Gesetz der großen Zahlen darbietet; dieses Gesetz fordere nämlich, daß die Zahl der Versuche im Vergleich zur Zahl der möglichen Fälle hinreichend groß sei, und daraus gehe hervor, daß jenes Gesetz auf Probleme keine Anwendung finden könne, in welchen die Menge der möglichen Fälle transfinit sei, weil, wie groß auch die Zahl der Versuche sein möge, sie doch immer unendlich klein im Vergleich zur Menge der Möglichkeiten wäre. Wie man sieht, stützt sich diese Argumentation auf denselben Gedanken, welcher der Brunsschen Formulierung des Gesetzes als einer gleichmäßigen Erschöpfung der möglichen *Fälle* zugrunde liegt. In Wirklichkeit handelt es sich in der Regel, bei geometrischen Wahrscheinlichkeiten immer, um Kategorien von Fällen, und dann braucht die Zahl der Versuche mit der Menge der Einzelfälle in keinem bestimmten Verhältnisse zu stehen.

88. Erfahrungsdaten aus Lotterieziehungen.

I. Die Untersuchung der Ziehungsergebnisse in der Prager und Brünner Lotterie, welche der Verfasser durchgeführt hat[1]) und die sich auf den Zeitraum 1754—1886, bzw. 1771—1886 erstrecken, hat nach jeder Richtung gute Übereinstimmung mit der Theorie ergeben.

Wenn man die Ziehungen nach ihrer Zusammensetzung aus ein- und zweiziffrigen Nummern sondert, so zerfallen sie in sechs Kategorien, je nachdem keine, eine, zwei, ... fünf einziffrige Nummern auftreten; die Wahrscheinlichkeiten dieser Kategorien sind in Nr. 38 gerechnet worden. Bezeichnet man mit s die Anzahl der Ziehungen, mit m_i die der Ziehungen einer Kategorie, so zeigen die Quotienten $\frac{m_i}{s}$ den Wahrscheinlichkeiten p_i gegenüber folgenden Verlauf:

Kategorie	p	Prag: $s = 2854$	Brünn: $s = 2703$
		$\frac{m_i}{s}$	$\frac{m_i}{s}$
0	0,58298	0,58655	0,57899
1	0,34070	0,32656	0,34591
2	0,06989	0,07919	0,06881
3	0,00619	0,00735	0,00629
4	0,00023	0,00035	0,00000
5	0,00000	0,00000	0,00000
	1,00000	1,00000	1,00000

Die Wahrscheinlichkeit, daß die Nummern einer Ziehung in der natürlichen *oder* in der umgekehrten Reihenfolge erscheinen, ist

1) Zum Gesetz der großen Zahlen. Prag 1889.

$$p = \frac{2}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = \frac{1}{60} = 0{,}01667;$$

die relative Häufigkeit $\frac{m}{s}$ derartiger Ziehungen war in

Prag: 0,01612, Brünn: 0,01739,

bei Vereinigung der Ergebnisse beider Orte

0,01674.

Die Wahrscheinlichkeit des Erscheinens einer einziffrigen Nummer ist für jeden Zug

$$p = 0{,}1;$$

die relative Häufigkeit einziffriger Nummern betrug in

Prag: 0,10168, Brünn: 0,10048,

in beiden Orten zusammen

0,10109.

Zu einer eingehenderen Untersuchung gibt die Tabelle der Erscheinungshäufigkeit der einzelnen Nummern Anlaß. Man kann nämlich das ganze Ziehungsmaterial — hier soll es für die Prager Ziehungen geschehen — als eine Folge von 90 gleich umfangreichen Beobachtungsreihen ansehen, angestellt über ein Ereignis der Wahrscheinlichkeit

$$p = \frac{1}{18},$$

und kann nun die Verhältniszahlen $\frac{m_i}{s}$ $(i = 1, 2, \ldots 90)$ auf ihre Verteilung um den Wert p prüfen; wir nehmen statt dessen die Prüfung an der Verteilung der Wiederholungszahlen m_i gegenüber der wahrscheinlichsten: $sp = 2754 \cdot \frac{1}{18} = 158$ selbst vor. Die Ergebnisse und die für die Rechnung erforderlichen Größen sind aus nachstehender Tabelle zu entnehmen.

Nummern	in der Anzahl	mit der Wiederholungs-zahl m	Abweichung $m - sp$	$(m - sp)^2$
6	1	138	— 20	400
39, 65	2	139	— 19	361
16, 41, 76, 87	4	142	— 16	256
2, 14, 56, 79, 86	5	143	— 15	225
18, 44, 47	3	144	— 14	196
72, 80	2	145	— 13	169
12	1	146	— 12	144
21, 53	2	147	— 11	121
70	1	149	— 9	81
24, 32, 55, 69	4	150	— 8	64
27, 64, 75	3	151	— 7	49
81	1	152	— 6	36
23, 29, 85	3	153	— 5	25
19, 35, 42, 74	4	154	— 4	16
7, 20, 59	3	155	— 3	9
13, 34, 40, 67, 88	5	156	— 2	4
19, 52, 68	3	157	— 1	1
17, 82	2	158	0	0
15, 90	2	159	1	1
58	1	160	2	4
8, 25, 36	3	161	3	9
22	1	162	4	16
33, 57	2	163	5	25
51	1	164	6	36
3, 43, 45, 48	4	165	7	49
10, 26, 66	3	166	8	64
1, 5, 60, 84	4	167	9	81
50, 62	2	168	10	100
9, 61, 63	3	170	12	144
54, 73	2	171	13	169
49, 71, 78	3	172	14	196
28	1	173	15	225
37	1	176	18	324
30, 46	2	177	19	361
89	1	178	20	400
31	1	179	21	441
38	1	184	26	676
4	1	185	27	729
77	1	186	28	784
83	1	189	31	961
	90	14270		13059

In Nr. 83 sind 150, 166 als wahrscheinliche Grenzen für die Wiederholungszahl einer Nummer in 2854 Ziehungen berechnet worden; in Wirklichkeit ist bei

$$4 + 3 + 1 + 3 + 4 + 3 + 5 + 3 + 2 + 2 + 1 + 3 + 1 + 2 + 1 + 4 + 3 = 45,$$

also genau bei der Hälfte der Nummern eine zwischen den genannten Grenzen (diese mit eingeschlossen) liegende Wiederholungszahl eingetroffen.

Für die mittlere und durchschnittliche Abweichung ergeben sich nach den Formeln (14), (17) die theoretischen Werte:

$$\sqrt{2854 \cdot \frac{1}{18} \cdot \frac{17}{18}} = 12{,}23,$$
$$\sqrt{\frac{2 \cdot 2854 \cdot 17}{18^2 \pi}} = 9{,}76;$$

andererseits berechnet sich aus der Summe der $(m - sp)^2$, welche 13059 beträgt, und aus der Summe der absoluten Werte der $(m - sp)$, welche 889 ausmacht, die mittlere und die durchschnittliche Abweichung:

$$\sqrt{\frac{13059}{90}} = 12{,}04,$$
$$\frac{889}{90} = 9{,}88,$$

also in sehr naher Übereinstimmung mit den theoretischen Werten.

II. G. Th. Fechner[1]) hat die Listen von zehn sächsischen Staatslotterien aus der Zeit von 1843—1852, je 32000 bis 34000 Nummern umfassend, in folgender Weise einer Prüfung unterzogen. Er teilte die gezogenen Nummern in der Reihenfolge, in der sie erschienen waren, in Serien zu je 3, dann zu je 10, 50, 100 Nummern und schied diese Serien nach dem absoluten Werte des Unterschiedes u zwischen den geraden und den ungeraden Nummern in Kategorien, bestimmte nach den Regeln der Wahrscheinlichkeitsrechnung die wahrscheinlichste Anzahl der in jede Kategorie fallenden Serien und verglich sie mit der wirklich beobachteten.

Im Sinne der in Nr. 86 eingeführten Bezeichnung bedeute s die Zahl der Nummern in einer Serie, s die Anzahl der Serien.

Für $s = 3$ sind 1 und 3 die möglichen Werte von $|u|$; der erstere stellt sich ein, wenn eine gerade und zwei ungerade Nummern, oder umgekehrt, gezogen werden; die Wahrscheinlichkeit hierfür ist

$$2\binom{3}{1}\frac{1}{2^3} = \frac{3}{4};$$

der zweite Wert tritt ein, wenn drei gerade oder drei ungerade Nummern erscheinen, wofür

$$2 \cdot \frac{1}{2^3} = \frac{1}{4}$$

1) Kollektivmaßlehre. Herausgegeben von G. F. Lipps (1897), p. 229 ff. — Wiewohl die gezogenen Nummern nicht zurückgelegt werden und daher die Zusammensetzung der Urne im Laufe der Ziehungen sich ändert, so ist doch von vornherein die Wahrscheinlichkeit für das Erscheinen einer geraden und ungeraden Nummer $\frac{1}{2}$ für jede einzelne Ziehung.

die Wahrscheinlichkeit ist. Da nun solcher Serien $z = 2000$ gebildet wurden, so wäre die wahrscheinlichste Verteilung auf die beiden Kategorien:

$$1500, \quad 500;$$

wirklich gezählt wurden

$$1494, \quad 506.$$

Bei $s = 10$ sind 0, 2, 4, 6, 8, 10 die möglichen Werte von $|u|$ und

$$\frac{10!}{5!\,5!}\,\frac{1}{2^{10}},\quad 2\,\frac{10!}{4!\,6!}\,\frac{1}{2^{10}},\quad 2\,\frac{10!}{3!\,7!}\,\frac{1}{2^{10}},\quad 2\,\frac{10!}{2!\,8!}\,\frac{1}{2^{10}},\quad 2\,\frac{10!}{1!\,9!}\,\frac{1}{2^{10}},\quad 2\cdot\frac{1}{2^{10}}$$

die bezüglichen Wahrscheinlichkeiten; durch ihre Multiplikation mit z ergibt sich die wahrscheinlichste Verteilung auf die einzelnen Kategorien.

Bei $s = 50$ gehen die Werte von $|u|$ von 0 bis 50, bei $s = 100$ von 0 bis 100; die Wahrscheinlichkeiten können bei dieser Größe von s nach der Näherungsformel

$$\frac{2}{\sqrt{2\pi s p q}}\,e^{-\frac{l^2}{2spq}}$$

gerechnet werden, in welcher $p = q = \frac{1}{2}$ und $l = \frac{|u|}{2}$ zu setzen ist; durch Multiplikation mit z ergibt sich wieder die wahrscheinlichste Verteilung der Serien auf die verschiedenen Werte von $|u|$, die nun mit der wirklich eingetroffenen zu vergleichen ist. Die Resultate sind in der nachfolgenden Tabelle zusammengestellt.

$\|u\|$	$s = 10$; $z = 5000$		$s = 50$; $z = 1000$		$s = 100$; $z = 600$	
	wahrscheinlichste Verteilung	beobachtete Verteilung	wahrscheinlichste Verteilung	beobachtete Verteilung	wahrscheinlichste Verteilung	beobachtete Verteilung
0	1230	1201	112	110	48	46
2	2051	2027	216	217	93,5	104
4	1172	1225	192	194	88	85
6	439	442	158	154	80	67
8	98	97	119,5	120	69,5	68
10	10	8	84	65	58	63
12	—	—	54	62	47	51
14	—	—	32	41	36	31
16			17	21	27	34
18			9	10	19	13
20			4	3	13	14
22			2	2	8,5	8
24			0,5	1	5,5	7
26			—	—	3	4
28			—	—	2	2
30					1	1
32					0,5	0
34					0,3	1
36					0,1	1
38					0,1	0
	5000	5000	1000	1000	600	600

89. Die Aufstellungen von K. Marbe. Als Beleg dafür, daß eine nicht ganz korrekte Anwendung der Theorie auf Erfahrungsdaten zu voreiligen Schlüssen über das Verhältnis der Wahrscheinlichkeitsrechnung zur Wirklichkeit und hierdurch zu Zweifeln an den Fundamenten der Theorie führen kann, seien die Untersuchungen von Marbe[1]) angeführt.

Marbe hat Ergebnisse des Roulettespiels[2]), die er in sehr beträchtlichem Umfange aus verschiedenen Spielorten (Monte Carlo, Spaa, Dinant) sich verschafft hatte, auf die Häufigkeit des Auftretens sogenannter „reiner" Gruppen oder Sequenzen, das sind Gruppen gleicher Ereignisse, geprüft. Bei der Bildung der Gruppen ist er, um möglichst große Versuchszahlen zu erzielen, nach einem Schema vorgegangen, das man als das Schema der *übergreifenden* Gruppen bezeichnen könnte. Numeriert man nämlich die Spielergebnisse, „rot" und „schwarz" (mit Auslassung von zéro), ihrer Zeitfolge nach, so werden die zweigliedrigen Gruppen nach dem Schema

$$(1, 2),\ (2, 3),\ (3, 4), \ldots,$$

die dreigliedrigen nach dem Schema

$$(1, 2, 3),\ (2, 3, 4),\ (3, 4, 5), \ldots.$$

usw. gebildet.

Die Anzahl der in dieser Weise aufgestellten r-gliedrigen Gruppen heiße N_r. Die Wahrscheinlichkeit einer r-gliedrigen reinen Gruppe ist $2\left(\frac{18}{37}\right)^r$; daher ist die erwartungsmäßige oder wahrscheinlichste Anzahl reiner Gruppen unter den N_r:

$$2\, N_r \left(\frac{18}{37}\right)^r,$$

mit der die beobachtete Anzahl solcher Gruppen zu vergleichen ist.

Diese Vergleichung, für $r = 1, 2, 3, \cdots$ ausgeführt, ließ nun *systematische* Abweichungen zwischen Theorie und Wirklichkeit zutage treten, die Marbe den Anlaß zu weitgehenden Schlußfolgerungen boten. Bevor auf diese eingegangen wird, sei hier eine solche Beobachtungsreihe (Marbes Tabelle V) als typisches Beispiel für alle angeführt.

---

1) Naturphilosophische Untersuchungen zur Wahrscheinlichkeitslehre. Leipzig 1899.

2) Beim Roulettespiel rotiert eine leicht drehbare horizontale Scheibe, deren Rand in 37 mit 0, 1, 2, . . . 36 bezeichnete und — mit Ausnahme von 0 — abwechselnd rot und schwarz gestrichene Fächer geteilt ist; gegen Ende der Rotation rollt aus einer konzentrischen Rinne ein Kügelchen in eines der Fächer und entscheidet das Spiel.

Gliederzahl der Gruppen (r)	Anzahl der beob. Gruppen (N_r)	Wahrscheinlichste Anzahl der reinen Gruppen $2\,N_r\left(\frac{18}{37}\right)^r$	Beobachtete Anzahl der reinen Gruppen	Verhältnis der beiden letzten Zahlen
1	11483	11173	11184	1·00
2	11467	5428	5461	1·00
3	11451	2637	2704	0·98
4	11435	1281	1348	0.95
5	11419	622	668	0·93
6	11403	302	317	0·95
7	11387	147	143	1·03
8	11371	71	63	1·13
9	11355	35	34	1·03
10	11339	17	7	2·43
11	11323	8	2	4·00
12	11307	4	0	∞
13	11291	2	0	∞
14	11275	1	0	∞

Als bleibende Erscheinung in allen bearbeiteten Beobachtungsreihen ergab sich, daß die hochgliedrigen Gruppen, die nach dem Maße ihrer Wahrscheinlichkeit noch hätten erwartet werden können, ausblieben, und daß schon vor dem gänzlichen Ausbleiben dieser Gruppen ein Zurückbleiben gegenüber der rechnungsmäßigen Erwartung eintrat. Marbe zog daraus den Schluß, daß reine Gruppen, die über eine gewisse Gliederzahl (bei der vorliegenden Materie etwa 12) hinausreichen, überhaupt nicht, wie groß auch die Gesamtgruppenzahl sein möge, eintreffen, und daß sie von einer gewissen Gliederzahl aufwärts notwendig hinter der Erwartung zurückbleiben. Um diese sehr wesentliche Abweichung der Wirklichkeit von den Angaben der Wahrscheinlichkeitsrechnung, aus der er sehr gewichtige Konsequenzen bezüglich der Grundlagen der letzteren zieht, verständlich zu machen, beruft sich Marbe zunächst auf D'Alembert, der in seinen Schriften[1]) wiederholt die Ansicht vertrat, daß reine Gruppen sich anders verhielten als gemischte Gruppen des gleichen Umfangs, der den ersteren von einer Gliederzahl an zwar logische Möglichkeit einräumt, sie aber physisch für unmöglich hält und damit indirekt die Unabhängigkeit der Fälle leugnet. Danach versucht Marbe die Erscheinung durch naturphilosophische Betrachtungen zu erklären, denen, abgesehen davon, daß sie sich auf den speziellen Vorgang beim Roulettespiel beziehen und daher des allgemeinen Charakters entbehren, eine Beweiskraft nicht zuerkannt werden kann.

Der wahre Grund der scheinbaren Anomalie liegt aber in der Art der Gruppenbildung, in dem Übergreifen der Gruppen. H. Grün-

1) Mélanges de littérature, d'histoire et de philosophie. Amsterdam 1767, p. 275 sq., 298.

baum[1]) hat hierüber eine eingehende Analyse durchgeführt und die Ergebnisse derselben auf neue Beobachtungsreihen angewendet, die sich von den Marbeschen eigentlich nur dadurch unterscheiden, daß die Grundwahrscheinlichkeit $\frac{1}{2}$ ist gegenüber $\frac{18}{37}$; denn der ungleiche Ursprung kann keinen Unterschied bedingen. Es sind Ziehungsresultate der ungarischen Klassenlotterie und Geburtsdaten von Schülern eines Gymnasiums, erstere nach ihrer Zeitfolge, letztere nach dem Alphabet geordnet, die in bezug auf das Auftreten gerader und ungerader Zahlen (im *Tages*datum; der 31. als Geburtstag wurde ausgelassen, um eine Begünstigung der ungeraden Zahlen hintanzuhalten) untersucht worden sind.

Grünbaum teilt die Beobachtungsreihen in *isolierte* reine Gruppen ein. Das Prinzip dieses Schemas ist folgendes. Bezeichnet 1 eine ungerade, 2 eine gerade Zahl, und ist

$$2\,1\,1\,2\,1\,2\,2\,2\,1\,1\,2\,\ldots.$$

der Anfang der Beobachtungsreihe, so beginnt die Gruppenbildung:

$$2;\ 11;\ 2;\ 1;\ 222;\ 11;\ 2\ldots.$$

Die Wahrscheinlichkeit einer r-gliedrigen isolierten reinen Gruppe ist $\frac{1}{2^{r+1}}$; denn sie setzt sich zusammen aus der Wahrscheinlichkeit der r-maligen Wiederholung einer und derselben Kategorie und der Wahrscheinlichkeit für das darauffolgende Auftreten der andern Kategorie. Hiernach bilden die erwartungsmäßigen Anzahlen der 1-, 2-, 3-.... gliedrigen isolierten Gruppen eine geometrische Reihe:

$$g_1,\ \frac{1}{2}g_1,\ \frac{1}{4}g_1,\ \ldots.$$

wenn g_1 die Anzahl der zu erwartenden eingliedrigen Gruppen bedeutet; die Gliedersummen dieser Gruppen sind

$$1g_1,\ 2\cdot\frac{1}{2}g_1,\ 3\cdot\frac{1}{4}g_1,\ \ldots.$$

Der Grenzwert der Summe der ersten Reihe ist $2g_1$, der Grenzwert der Summe der zweiten $4g_1$, folglich das Grenzverhältnis der Zahl der isolierten Gruppen zur Anzahl der Einzelglieder $\frac{1}{2}$.

Man hat also für eine sehr große Anzahl N von Beobachtungen die folgende Verteilung als die wahrscheinlichste zu erwarten:

$$\text{Gesamtzahl der isolierten Gruppen: } \frac{N}{2};$$

$$g_1 = \frac{N}{4},\ g_2 = \frac{N}{8},\ \ldots g_r = \frac{N}{2^{r+1}},\ \ldots$$

1) Isolierte und reine Gruppen und die Marbesche Zahl „p". Würzburg, 1904.

Nun gibt eine r-gliedrige *isolierte* Gruppe

$$\begin{aligned} r &\cdots \text{1-gliedrige} \\ r-1 &\cdots \text{2-gliedrige} \\ r-2 &\cdots \text{3-gliedrige} \\ &\ldots\ldots\ldots \\ 1 &\cdots r\text{-gliedrige} \end{aligned}$$

reine Gruppen im Marbeschen Sinne; umgekehrt gehen r-gliedrige *übergreifende* Gruppen hervor:

je 1 aus den r-gliedrigen,
„ 2 „ „ $\overline{r+1}$-gliedrigen,
„ 3 „ „ $\overline{r+2}$-gliedrigen

isolierten Gruppen, so daß sich die erwartungsmäßige Anzahl γ_r der übergreifenden r-gliedrigen Gruppen aus den isolierten Gruppen nach folgendem Gesetz ergibt:

$$\gamma_r = 1 \cdot g_r + 2g_{r+1} + 3g_{r+2} + \cdots;$$

dies gäbe bei *unbegrenzter* Versuchszahl

$$4g_r = 4\frac{N}{2^{r+1}}, = 2N\left(\frac{1}{2}\right)^r,$$

und dies stimmt mit der Anzahl der nach Marbes Schema zu erwartenden Gruppen überein, weil bei sehr großem N zwischen N_r und N ein im Verhältnis zu den Zahlen selbst nur unerheblicher Unterschied besteht. *Bei begrenzter* Versuchszahl ist aber die Reihe nur so weit zu führen, als isolierte Gruppen vermöge ihrer Wahrscheinlichkeit überhaupt noch zu erwarten sind; bezeichnet man diese höchste Gliederzahl mit p, so ist statt γ_r zu erwarten:

$$\gamma_r(p) = 1g_r + 2g_{r+1} + \cdots + (p - r + 1)g_p.$$

Indem sich nun die Marbesche Untersuchung auf die Zahlen γ_r statt, wie es richtig wäre, auf die Zahlen $\gamma_r(p)$ stützt, *überschätzt* sie die erwartungsmäßigen Zahlen der reinen Gruppen, und diese Überschätzung macht sich bei den hoch zusammengesetzten Gruppen am stärksten geltend. Ein numerisches Beispiel wird dies am besten klar machen.

Angenommen, die Beobachtungsreihe umfasse $N = 2^n$ Zahlen; dann sind an isolierten Gruppen zu erwarten:

$$\begin{aligned} 2^{n-1} &\cdots \text{überhaupt,} \\ 2^{n-2} &\cdots \text{1-gliedrige,} \\ 2^{n-3} &\cdots \text{2-gliedrige,} \\ &\ldots\ldots\ldots \\ 2^{n-n} = 1 &\cdots \overline{n-1}\text{-gliedrige;} \end{aligned}$$

von da ab werden die Anzahlen zu erwartender Gruppen $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}\cdots$, also kleiner als 1, d. h. Gruppen höherer Zusammensetzung sind überhaupt nicht zu erwarten. Marbe aber erwartet statt der *einen* $\overline{n-1}$-gliedrigen Gruppe deren

$$1 + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} + \cdots = 4;$$

statt der *vier* $\overline{n-2}$-gliedrigen Gruppen deren

$$1 \cdot 2 + 2 \cdot 1 + 3 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + \cdots = 8;$$

statt der *elf* $\overline{n-3}$-gliedrigen Gruppen deren

$$1 \cdot 4 + 2 \cdot 2 + 3 \cdot 1 + 4 \cdot \frac{1}{2} + \cdots = 16;$$

statt der *vierundzwanzig* $\overline{n-4}$-gliedrigen Gruppen deren 32 usw. Die Verhältniszahlen $\frac{4}{1}, \frac{8}{4}, \frac{16}{11}, \frac{32}{24}, \cdots$ werden immer kleiner und rücken an 1 heran, die Überschätzung nimmt also gegen die Gruppen niedriger Zusammensetzung hin immer mehr ab. Aber noch mehr; während n-gliedrige isolierte Gruppen nach obiger Rechnung in der Anzahl $\frac{1}{2}$, also streng genommen nicht mehr zu erwarten sind, erwartet sie Marbe in der Anzahl

$$1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \cdots = 2, \text{ usw.}$$

Nach dem Schema der übergreifenden Gruppen				Nach dem Schema der isolierten Gruppen			
Reine Gruppen, Gliederzahl	Wahrscheinlichste Anzahl	Beobachtete Anzahl	Verhältnis beider Anzahlen	Isolierte Gruppen, Gliederzahl	Wahrscheinlichste Anzahl	Beobachtete Anzahl	Verhältnis beider Anzahlen
1	1500	1500	1·00	überhaupt	750	782	0·96
2	750	718	1·04	1	375	407	0·92
3	375	343	1·09	2	187·5	188	1·00
4	187	156	1·19	3	93·7	101	0·93
5	94	70	1·34	4	46·8	45	1·04
6	47	29	1·62	5	23·4	26	0·90
7	23	14	1·64	6	11·7	8	1·46
8	12	7	1·71	7	5·8	3	1·93
9	6	3	2·00	8	2·9	2	1·45
10	3	1	3·00	9	1·4	1	1·40
11	1	0	∞	10	0·7	1	0·70
12	1	0	∞	11	0·3	0	—

Die vorstehende Tabelle zeigt eine der von Grünbaum benützten Beobachtungsreihen in beiden Bearbeitungen. Man ersieht aus der Bearbeitung nach dem Schema der übergreifenden Gruppen alle die

von **Marbe** konstatierten Anomalien, die in völlig befriedigender Weise verschwinden, wenn die korrekte Methode der isolierten Gruppen zur Anwendung gebracht wird.[1])

90. Würfelversuche. Wohl die umfangreichsten Versuchsreihen hat der Züricher Astronom R. **Wolf** ausgeführt.[2]) Eine davon, mit der größten Versuchszahl, aus dem Jahre 1850 stammend, bestand darin, daß mit zwei gewöhnlichen Würfeln 1000-mal so lange gewürfelt wurde, bis jeder der 21 möglichen Würfe (die Würfel wurden nicht unterschieden) wenigstens einmal zum Vorschein gekommen war; es waren dazu 97899 Würfe notwendig, die dann auf 100000 ergänzt worden sind; alle Würfe wurden notiert. Aus der Fülle der Resultate sei nur angeführt, daß sich die relative Häufigkeit eines unpaaren Wurfs

aus 100 Würfen mit 0·88,
„ 1000 „ „ 0·836,
„ 10000 „ „ 0·8351,
„ 100000 „ „ 0·83533

ergab, während die Wahrscheinlichkeit eines solchen Wurfes $\frac{5}{6} = 0{\cdot}83333\ldots$ beträgt.

Ausführlicher soll hier auf eine jüngere Versuchsreihe[3]) **Wolfs** eingegangen werden, weil ihre Ergebnisse geeignet sind, manches von den vorausgehenden allgemeinen Erörterungen zu illustrieren.

Mit zwei Würfeln, deren einer rot gebeizt war, so daß alle möglichen 36 Verbindungen der Würfelseiten unterschieden werden konnten, sind 20000 Würfe gemacht und die erschienenen Kombinationen verzeichnet worden. Das Resultat der Zählung war folgendes:

		Weißer Würfel					
	Nr.	1	2	3	4	5	6
Roter Würfel	1	547	587	500	462	621	690
	2	609	655	497	535	651	684
	3	514	540	468	438	587	629
	4	462	507	414	413	509	611
	5	551	562	499	506	658	672
	6	563	598	519	487	609	646

1) Eine allgemeine Theorie der Untersuchung von beobachteten Reihen zufälliger Ereignisse nach übergreifenden Gruppen, die sich weittragender Methoden bedient, hat H. **Bruns** gegeben: Das Gruppenschema für zufällige Ereignisse. Abhandl. d. Leipz. Ges. d. Wissensch., XXIX. Bd. (1906), p. 579—628.

2) Schriften der Naturforsch. Gesellsch. in Bern 1849 bis 1851, 1853; Naturforsch. Gesellsch. in Zürich, 38 (1893).

3) Drei Mitteilungen über neue Würfelversuche. Naturforsch. Gesellsch. in Zürich, 26, 27 (1881—1883).

Gleiche Möglichkeit der Fälle vorausgesetzt, ist $\frac{1}{36}$ die Wahrscheinlichkeit jeder einzelnen Kombination und $\frac{1}{36} \cdot 20000 = 556$ ihre wahrscheinlichste Wiederholungszahl. Die Abweichungen von dieser Zahl und ihre Quadrate sind aus der folgenden Zusammenstellung zu ersehen.

$s = 20000;\ p = \frac{1}{36}$		
m	$\vert m - sp \vert$	$(m - sp)^2$
413	143	20449
414	142	19884
438	118	13924
462	94	8836
462	94	8836
468	88	7744
487	69	4761
497	59	3481
499	57	3249
500	56	3136
506	50	2500
507	49	2401
509	47	2209
514	42	1764
519	37	1369
535	21	441
540	16	256
547	9	81
551	5	25
562	6	36
563	7	49
587	31	961
587	31	961
598	42	1764
609	53	2809
609	53	2809
611	55	3025
621	65	4225
629	73	5329
646	90	8100
651	95	9025
655	99	9801
658	102	10404
672	116	13456
684	128	16384
690	134	17956
20000	2376	212440

Die wahrscheinliche Abweichung

$$0{,}476936 \sqrt{2 \cdot 20000 \cdot \frac{1}{36} \cdot \frac{35}{36}} = 13, \cdots$$

führt zu den Grenzen 543, 569, zwischen welchen am wahrscheinlichsten die Hälfte der m, d. i. 18 derselben zu erwarten wären; in Wirklichkeit liegen aber nur 4 zwischen diesen Grenzen (547, 551, 562, 563).

Der theoretische Wert der mittleren Abweichung:

$$\sqrt{20000 \cdot \frac{1}{36} \cdot \frac{35}{36}} = 23{,}2$$

ergibt sich wesentlich verschieden von dem aus den wirklichen Abweichungen berechneten:

$$\sqrt{\frac{212440}{36}} = 76{,}8.$$

Ähnlich verhält es sich mit der durchschnittlichen Abweichung, deren theoretischer Wert sich mit 18,54 ergibt, während sie sich aus der Summe der $|m - sp|$ mit 66 berechnet.

Die Beobachtungsreihe zeigt also in jeder Beziehung so erhebliche Widersprüche gegenüber der Theorie, daß man mit Sicherheit den Schluß ziehen darf, es seien die Bedingungen, welche der wahrscheinlichkeitstheoretischen Behandlung zugrunde liegen, bei den Versuchen nicht erfüllt gewesen. Da nun die Art ihrer Ausführung eine Begünstigung einzelner Fälle ausschließt, so kann der Grund nur in Unregelmäßigkeiten der Würfel liegen, welche bewirkten, daß die Würfelseiten nicht als gleichmögliche Fälle sich darstellen. Dieser Umstand verrät sich schon durch den systematischen Charakter der Abweichungen, den ein Blick auf die erste Tafel deutlich erkennen läßt: die Kombinationen, welche die Seite 4 des einen wie des andern Würfels enthalten, sind weniger zahlreich vertreten, während die Kombinationen mit 6 auffällig häufig erschienen sind.[1])

91. Erfahrungen betreffend geometrische Materien. Das Nadelproblem hat zu wiederholten Versuchen Anlaß gegeben; wenn man in der dafür geltenden Formel (Nr. 53) $p = \frac{2c}{\pi a}$ für p die aus den Versuchen resultierende relative Häufigkeit der Fälle, wo die Nadel eine der Parallellinien kreuzt, einsetzt, so ergibt sich eine empirische Bestimmung für π; gerade dieser Umstand hat anregend gewirkt.

Die erste Versuchsreihe hat R. Wolf 1850 ausgeführt; die Parallelen waren im Abstande von $2a = 45$ mm gezogen, das aus

---

1) R. Wolf hat unter Hinzuziehung einer Hypothese über die Abhängigkeit der Wahrscheinlichkeit der Würfelseiten von der Lage des Schwerpunktes und weiterer Versuchsreihen folgende Wahrscheinlichkeiten ermittelt:

Nr.	1	2	3	4	5	6
Weißer Würfel	0,16230	0,17245	0,14485	0,14205	0,18175	0,19660
Roter Würfel	0,17035	0,18155	0,15880	0,14580	0,17240	0,17110

einer Stricknadel herausgebrochene Versuchsstück hatte die Länge $2c = 36$ mm. Wolf warf die Nadel in Serien zu je 100 und erteilte der Tafel nach jeder Serie eine kleine Drehung; im ganzen kam die Nadel in 2532 von den $50 \cdot 100 = 5000$ Würfen über eine der Parallelen zu liegen; mit dieser relativen Häufigkeit ergibt die Formel für π den Wert 3,1596.

Von späteren Versuchsreihen seien angeführt:[1])

M. A. Smith (1855) aus 3204 Würfen ··· 3,1553
Kapt. Fox (1894) aus 1120 „ ··· 3,1419
M. Lazzarini (1901) aus 3408 „ ··· 3,1415929.

R. Lämmel[2]) hat Versuche zum Bertrandschen Paradoxon (Nr. 71) ausgeführt; doch gewinnt man aus seinen Angaben keine klare Vorstellung über den eingeschlagenen Vorgang. Es scheint, daß es sich um den Modus 4) gehandelt habe, für den die Theorie die Wahrscheinlichkeit 0,609 ergibt. Eine Serie von $5 \cdot 100$ Sehnen (in Kreisen von 15 cm Radius) lieferte die relative Häufigkeit 0,588, eine zweite von $3 \cdot 100$ Sehnen den Wert 0,644; die Zusammenfassung beider Serien führte zu 0,609.[3])

§ 2. Das Theorem von Poisson.

92. Entwicklung der Fragestellung. Einen wichtigen Gedanken hat Poisson[4]) in jenen Teil der Wahrscheinlichkeitstheorie eingeführt, welcher sich mit der Erwartungsbildung in bezug auf die Ergebnisse wiederholter Versuche beschäftigt, indem er von der Voraussetzung konstant bleibender Wahrscheinlichkeiten abgehend annahm, daß sich die Wahrscheinlichkeiten der in Frage kommenden Ereignisse E, $\bar{E}$ im Laufe der Versuche ändern. Es lassen sich in der Tat leicht Vorgänge konstruieren, bei welchen dies zutrifft; sobald man den Boden der Anwendungen betritt, scheint die Variabilität der Wahrscheinlichkeiten die Regel zu bilden.

Das Problem, mit dem wir uns hier zunächst beschäftigen, besteht in folgendem: Es werden s Versuche oder Beobachtungen über die beiden entgegengesetzten Ereignisse E, $\bar{E}$ angestellt, denen im λ-ten

---

1) Periodico di matematica, 1901, p. 140—143.

2) l. c., p. 38—39.

3) Ich ließ kürzlich von meinen Hörern Versuche zum Modus 6) anstellen, dem die Wahrscheinlichkeit 0,746 entspricht. Von 2060 Sehnen fielen 1457 länger aus als die Dreiecksseite, woraus sich die relative Häufigkeit solcher Sehnen mit 0.7073 ergibt. Die Versuchsblätter zeigten, daß es nicht leicht sei, den Prozeß so zu führen, daß eine angenähert gleichmäßige Verteilung der Punkte Platz greift.

4) Recherches sur la probabilité etc. Paris 1837; deutsch von H. Schnuse, Braunschweig 1841, 4. Kapitel. — Außerdem Abhandlungen in den Compt. rend. 1, 2 (1835, 1836).

Versuche ($\lambda = 1, 2, \cdots s$) die Wahrscheinlichkeiten p_λ, $q_\lambda = 1 - p_\lambda$ zukommen. Gefragt wird nach der Wahrscheinlichkeit, daß sich die Ereignisse in einer bestimmten Anzahl von Malen wiederholen; nach derjenigen Kombination, welche unter allen die größte Wahrscheinlichkeit besitzt; endlich nach der Wahrscheinlichkeit, daß die Wiederholungszahlen m, n von E, $\bar{E}$ oder die Verhältnisse $\frac{m}{s}$, $\frac{n}{s}$, welche die relative Häufigkeit des Erscheinens ausdrücken, vorgegebene Grenzen nicht überschreiten.

Zur schematischen Darstellung der Sachlage denke man sich s Urnen U_λ ($\lambda = 1, 2, \cdots s$), welche mit weißen und schwarzen Kugeln in solchem Mengenverhältnis gefüllt sind, daß die Urne U_λ dem Erscheinen einer weißen Kugel die Wahrscheinlichkeit p_λ, dem Erscheinen einer schwarzen Kugel die Wahrscheinlichkeit q_λ verleiht; aus jeder Urne wird eine Kugel gezogen. Wiewohl die Reihenfolge, in welcher die Urnen darankommen, für den Erfolg gleichgültig ist, stelle man sich, um mit der obigen Formulierung im Einklange zu bleiben, vor, daß die λ-te Ziehung aus der Urne U_λ geschehe.

Bei kleinem s kann die Beantwortung der oben gekennzeichneten Fragen auf dem Wege der direkten Rechnung geschehen. Beschwerlich wird dieser Weg aber schon bei einigermaßen großem s (etwa von der Ordnung einiger Zehner), und ganz unverwendbar bei sehr großem s. Für diesen Fall hat nun Poisson Näherungsformeln entwickelt mit Hilfe einer Analyse, welche in ihrer Grundlage auf Laplace[1]) zurückführt.

93. Ableitung des Theorems. Die Entwicklung des Produktes

$$\prod_1^s (p_\lambda u + q_\lambda v)$$

nach Potenzen von u und v befolgt solche arithmetische Regeln, daß der Koeffizient von $u^m v^n$ die Wahrscheinlichkeit für das m-malige Eintreffen von E und das n-malige Eintreffen von $\bar{E}$ bedeutet. Bezeichet man diese Wahrscheinlichkeit mit $U_{m,n}$ und ersetzt u durch e^{xi}, v durch e^{-xi}, unter i die imaginäre Einheit verstanden, so wird

$$X = \prod_1^s (p_\lambda e^{xi} + q_\lambda e^{-xi}) = \sum U_{m,n} e^{(m-n)xi}.$$

Durch Multiplikaiion mit $e^{-(m-n)xi}$ entfällt der exponentielle Faktor bei $U_{m,n}$ und nur bei diesem Gliede. Integriert man hierauf in bezug auf x zwischen den Grenzen $-\pi$, π, so fallen rechter Hand

1) Théorie analyt., art. 8.

alle Glieder bis auf das mit $U_{m,n}$ behaftete aus, weil für jedes ganzzahlige r (mit Ausschluß der Null)

$$\int_{-\pi}^{\pi} e^{rxi}\,dx = 0$$

ist, und es ergibt sich für $U_{m,n}$ die Darstellung:

$$U_{m,n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} X e^{-(m-n)xi}\,dx\,.$$

Nun ist

$$X = \prod_1^s \{(p_\lambda + q_\lambda)\cos x + i\,(p_\lambda - q_\lambda)\sin x\}$$

$$= \prod_1^s \{\cos x + i\,(p_\lambda - q_\lambda)\sin x\}$$

$$= \prod_1^s \{\varrho_\lambda(\cos\varphi_\lambda + i\sin\varphi_\lambda)\}\,,$$

wenn

$$\varrho_\lambda = \sqrt{\cos^2 x + (p_\lambda - q_\lambda)^2\sin^2 x} = \sqrt{1 - 4p_\lambda q_\lambda \sin^2 x} \tag{1}$$

und

$$\operatorname{tg}\varphi_\lambda = (p_\lambda - q_\lambda)\operatorname{tg} x \tag{2}$$

gesetzt wird; mit den Abkürzungen

$$\prod_1^s \varrho_\lambda = \mathrm{R}, \qquad \sum_1^s \varphi_\lambda = \Phi \tag{3}$$

wird also

$$X = R\,(\cos\Phi + i\sin\Phi)$$

und

$$U_{m,n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} R\{\cos(\Phi - \overline{m-n}\,x) + i\sin(\Phi - \overline{m-n}\,x)\}\,dx$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} R\cos(\Phi - \overline{m-n}\,x)\,dx\,,$$

wenn man beachtet, daß R vermöge (1) eine gerade, Φ hingegen wegen (2) eine ungerade Funktion von x ist. Aus denselben Gründen kann auch

$$U_{m,n} = \frac{1}{\pi}\int\limits_0^{\pi} R\cos(\Phi - \overline{m-n}x)\,dx$$

geschrieben werden.

Zerlegt man das Integrationsgebiet in die Teile $\left(0, \frac{\pi}{2}\right)$, $\left(\frac{\pi}{2}, \pi\right)$ und substituiert in dem zweiten Teile $x = \pi - \xi$, so geht Φ über in $(m+n)\pi - \Phi$, während R keine Änderung erfährt; das zweite Teilintegral lautet dann

$$\int\limits_0^{\frac{\pi}{2}} R\cos(\overline{m+n}\pi - \Phi + \overline{m-n}\xi - \overline{m-n}\pi)\,d\xi,$$

und dies ist gleich

$$\int\limits_0^{\frac{\pi}{2}} R\cos(2n\pi - \Phi + \overline{m-n}\xi)\,d\xi = \int\limits_0^{\frac{\pi}{2}} R\cos(\Phi - \overline{m-n}\xi)\,d\xi,$$

stimmt also dem Werte nach mit dem ersten Teilintegral überein, so daß

$$U_{m,n} = \frac{2}{\pi}\int\limits_0^{\frac{\pi}{2}} R\cos(\Phi - \overline{m-n}x)\,dx \tag{4}$$

ist.

Unter der Voraussetzung, daß kein p_λ und q_λ der Null oder der Einheit sehr nahe liegt, kommt kein Faktor von

$$R = \prod_1^s \sqrt{1 - 4p_\lambda q_\lambda \sin^2 x}$$

der Einheit nahe, und es wird R um so kleiner, je weiter sich x von der unteren Integralgrenze 0 entfernt; belangreiche Werte nimmt es nur in der Nähe dieser Grenze an und kann hier annähernd durch

$$\prod_1^s (1 - 2p_\lambda q_\lambda x^2),$$

der Logarithmus von R durch

$$-x^2 \sum 2p_\lambda q_\lambda$$

ersetzt werden; mit der abkürzenden Bezeichnung

$$k = \sqrt{\frac{2\sum p_\lambda q_\lambda}{s}}\ ^{1)} \tag{5}$$

wird also näherungsweise

$$R = e^{-s k^2 x^2}. \tag{6}$$

Weiter folgt aus

$$\varrho_\lambda \sin \varphi_\lambda = (p_\lambda - q_\lambda) \sin x$$

mit Benützung des Wertes von ϱ_λ aus (1) und bei Beschränkung auf kleine Werte von x

$$\sin \varphi_\lambda = (p_\lambda - q_\lambda)(1 + 2 p_\lambda q_\lambda x^2)\left(x - \frac{x^3}{6} + \cdots\right)$$

$$= (p_\lambda - q_\lambda)\, x + (p_\lambda - q_\lambda)\left(2 p_\lambda q_\lambda - \frac{1}{6}\right) x^3 + \cdots,$$

daraus durch Inversion

$$\varphi_\lambda = (p_\lambda - q_\lambda)\, x + \frac{4}{3}(p_\lambda - q_\lambda)\, p_\lambda q_\lambda x^3;$$

demzufolge ist, wenn man die abkürzenden Bezeichnungen:

$$\frac{\sum p_\lambda}{s} = \frac{p_1 + p_2 + \cdots + p_s}{s} = p, \quad \frac{\sum q_\lambda}{s} = \frac{q_1 + q_2 + \cdots + q_s}{s} = q \tag{7}$$

$$\frac{4}{3}\frac{\sum (p_\lambda - q_\lambda)\, p_\lambda q_\lambda}{s} = h \tag{8}$$

einführt,

$$\Phi = s(p - q)\, x + s h x^3. \tag{9}$$

Mit diesen Näherungswerten von R und Φ stellt sich jetzt $U_{m,n}$ wie folgt dar:

$$U_{m,n} = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} e^{-s k^2 x^2} \cos\left[\{s(p-q) - (m-n)\}\, x + s h x^3\right] dx.$$

Der Koeffizient von x kann umgeformt werden in

$$s\left\{p - \frac{m}{s} - \left(q - \frac{n}{s}\right)\right\} = s g, \tag{10}$$

---

1) Für den numerischen Wert von k läßt sich eine obere Grenze angeben. Da nämlich $\frac{1}{4}$ der größtmögliche Wert des Produktes $p_\lambda q_\lambda$ ist, so kann $\sum p_\lambda q_\lambda$ den Wert $\frac{s}{4}$ nicht überschreiten; infolgedessen ist

$$k \leqq \sqrt{\frac{1}{2}} = 0{,}7071 \cdots,$$

wobei, vermöge $p + q = 1$ und $m + n = s$, g den doppelten Unterschied $p - \frac{m}{s}$ $\left(\text{oder auch den doppelten Unterschied } \frac{n}{s} - q\right)$ bedeutet. Wird hierauf der Kosinus entwickelt, hierbei aber über die dritte Potenz von x nicht hinausgegangen, so verwandelt sich der obige Ausdruck in:

$$U_{m,n} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} e^{-sk^2x^2} \{\cos(sgx) - shx^3 \sin(sgx)\} \, dx$$

und durch die Substitution

$$x\sqrt{s} = z$$

weiter in:

$$U_{m,n} = \frac{2}{\pi\sqrt{s}} \int_0^{\frac{\pi}{2}\sqrt{s}} e^{-k^2z^2} \left\{\cos\left(gz\sqrt{s}\right) - \frac{hz^3}{\sqrt{s}} \sin\left(gz\sqrt{s}\right)\right\} dz.$$

Wegen der mit wachsendem z außerordentlich rasch fallenden Exponentialfunktion wird der Wert der Funktion unter dem Integralzeichen schon innerhalb des Integrationsintervalls sehr klein und außerhalb desselben so geringfügig, daß es auf den Wert von $U_{m,n}$ keinen merkbaren Einfluß übt, wenn man die obere Grenze durch ∞ ersetzt.

Nun ist[1])

$$\int_0^{\infty} e^{-k^2z^2} \cos\left(gz\sqrt{s}\right) dz = \frac{\sqrt{\pi}}{2k} e^{-\frac{sg^2}{4k^2}};$$

daraus ergibt sich durch Differentiation nach g:

$$\int_0^{\infty} e^{-k^2z^2} z\sqrt{s} \sin\left(gz\sqrt{s}\right) dz = \frac{sg\sqrt{\pi}}{4k^3} e^{-\frac{sg^2}{4k}},$$

weiter durch Differentiation in bezug auf k:

$$\int_0^{\infty} e^{-k^2z^2} 2kz^3\sqrt{s} \sin\left(gz\sqrt{s}\right) dz = \left(\frac{3sg\sqrt{\pi}}{4k^4} - \frac{sg^3\sqrt{\pi}}{8k^6}\right) e^{-\frac{sg^2}{4k^2}},$$

1) Man kommt zu dem Werte J dieses Integrals, indem man es nach g differentiiert, das Integral $\frac{dJ}{dg}$ durch partielle Integration entwickelt und die so entstandene Differentialgleichung integriert; zu dem Werte der Integrationskonstanten führt die Annahme $g = 0$.

woraus sich

$$\int_0^\infty e^{-k^2 z^2} \frac{hz^3}{\sqrt{s}} \sin(gz\sqrt{s})\, dz = -\frac{gh\sqrt{\pi}}{8k^5}\left(3 - \frac{sg^2}{2k^2}\right) e^{-\frac{sg^2}{4k^2}}$$

berechnet.

Nach Einführung dieser Integralwerte in den letzten Ausdruck für $U_{m,n}$ wird, wenn man noch

$$\frac{g\sqrt{s}}{2k} = \theta \tag{11}$$

setzt,

$$U_{m,n} = \frac{1}{k\sqrt{\pi s}}\left[1 - \frac{h\theta}{2k^3\sqrt{s}}(3 - 2\theta^2)\right] e^{-\theta^2}. \tag{12}$$

Wenn man auf die Bedeutung von g zurückgeht, so drückt $U_{m,n}$ die Wahrscheinlichkeit aus, daß die Ereignisse E, $\bar{E}$ in den Wiederholungszahlen

$$m = sp - k\theta\sqrt{s}, \quad n = sq + k\theta\sqrt{s}$$

erscheinen.

Für die Wiederholungszahlen

$$m' = sp + k\theta\sqrt{s}, \quad n' = sq - k\theta\sqrt{s}$$

besteht dementsprechend die Wahrscheinlichkeit

$$U_{m',n'} = \frac{1}{k\sqrt{\pi s}}\left[1 + \frac{h\theta}{2k^3\sqrt{s}}(3 - 2\theta^2)\right] e^{-\theta^2}. \tag{13}$$

Hiernach ist die Wahrscheinlichkeit, daß sich die Wiederholungszahl des Ereignisses E von der Zahl sp um $k\theta\sqrt{s}$ nach auf- oder abwärts unterscheide,

$$U_{m,n} + U_{m',n'} = \frac{2}{k\sqrt{\pi s}} e^{-\theta^2},$$

und dies ist am größten für $\theta = 0$.

Daraus ergibt sich als erstes Resultat:

„Die wahrscheinlichste unter allen Kombinationen ist diejenige, in welcher sich die Wiederholungszahlen der Ereignisse E und $\bar{E}$ so verhalten wie ihre durchschnittlichen Wahrscheinlichkeiten p, q im Laufe der Versuche.“

Denn nach den Formeln (7) sind p, q die arithmetischen Mittel der Wahrscheinlichkeiten p_λ, q_λ und ergänzen sich so wie diese zur Einheit.

Setzt man

$$k\theta\sqrt{s} = \xi$$

und denkt sich unter ξ eine ganze Zahl, so wird die Wahrscheinlich-

keit P, daß die Wiederholungszahl m von E zwischen die Grenzen $sp - l$ und $sp + l$ falle, ausgedrückt sein durch

$$\sum_0^l \frac{2}{k\sqrt{\pi s}} e^{-\frac{\xi^2}{k^2 s}} - \frac{1}{k\sqrt{\pi s}};$$

der negative Teil rührt daher, daß in der Summe die Wahrscheinlichkeit der Abweichung 0 doppelt gezählt ist. Verwandelt man die Summe nach der in Nr. 76 angegebenen Methode in ein Integral, so wird

$$P = \frac{2}{k\sqrt{\pi s}} \int_0^l e^{-\frac{\xi^2}{k^2 s}} d\xi + \frac{e^{-\frac{l^2}{k^2 s}}}{k\sqrt{\pi s}},$$

oder einfacher, wenn man

$$\frac{\xi}{k\sqrt{s}} = t, \qquad \frac{l}{k\sqrt{s}} = \gamma$$

setzt,

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt + \frac{e^{-\gamma^2}}{k\sqrt{\pi s}}. \tag{14}$$

Die zweite Aussage, die mit der vorigen zusammen *Poissons Theorem* ausmacht, lautet demnach wie folgt:

„Es ist mit der Wahrscheinlichkeit

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt + \frac{e^{-\gamma^2}}{k\sqrt{\pi s}}$$

zu erwarten, daß die Wiederholungszahl m des Ereignisses E zwischen die Grenzen

$$sp \mp \gamma k \sqrt{s},$$

die relative Häufigkeit $\frac{m}{s}$ dieses Ereignisses also zwischen die Grenzen

$$p \mp \frac{\gamma k}{\sqrt{s}}$$

fallen werde."

Die wichtigste Folgerung aus diesem Satze geht dahin, daß durch entsprechende Vermehrung der Versuche die Grenzen, innerhalb welcher das Verhältnis $\frac{m}{s}$ mit einer der Einheit beliebig nahen Wahrscheinlichkeit zu erwarten ist, beliebig eng gezogen werden können. Bei der Begründung dieser Schlußfolgerung ist auf den im Gange des Beweises hervorgehobenen Umstand Rücksicht zu nehmen, daß k un-

abhängig von der Größe der Versuchszahl an eine obere Grenze $\left(\sqrt{\frac{1}{2}}\right)$ gebunden ist.

Das Bernoullische Theorem ist als besonderer Fall in dem Poissonschen enthalten; denn, werden die Wahrscheinlichkeiten p_λ untereinander gleich und gleich p, so sind auch die q_λ einander gleich und gleich q;

$$k = \sqrt{\frac{2\sum p_\lambda q_\lambda}{s}}$$

geht über in $\sqrt{2pq}$, und die Ausdrücke für P, die Grenzen von m und $\frac{m}{s}$ verwandeln sich in diejenigen, welche das Bernoullische Theorem anführt.

Eine Bemerkung möge schon an dieser Stelle Platz finden. In bezug auf das wahrscheinlichste Ergebnis der s Versuche ist es ebenso, als ob sie bei konstant bleibender Wahrscheinlichkeit $p = \frac{\sum p_\lambda}{s}$ ausgeführt worden wären; auf andere Beziehungen darf aber diese Analogie nicht übertragen werden.

94. Beispiel XLVII. *Es liegen drei Urnen A_1, A_2, A_3 vor; die erste enthält 2 weiße und 3 schwarze, die zweite 3 weiße und 2 schwarze, die dritte 1 weiße und 4 schwarze Kugeln. Man macht aus A_1 200, aus A_2 400, aus A_3 600 Ziehungen, wobei die gezogene Kugel jedesmal wieder zurückgelegt wird. Welches ist das wahrscheinlichste Resultat dieser Ziehungen und welches die wahrscheinliche Abweichung von diesem?*

Die durschschnittlichen Wahrscheinlichkeiten für die weiße und schwarze Kugel sind

$$p = \frac{\frac{2}{5} + 2 \cdot \frac{3}{5} + 3 \cdot \frac{1}{5}}{6} = \frac{11}{30}, \qquad q = \frac{\frac{3}{5} + 2 \cdot \frac{2}{5} + 3 \cdot \frac{4}{5}}{6} = \frac{19}{30}.$$

Mit diesen berechnen sich die wahrscheinlichsten Wiederholungszahlen

$$sp = 1200 \cdot \frac{11}{30} = 440, \qquad sq = 1200 \cdot \frac{19}{30} = 760.$$

Ferner ist

$$k = \sqrt{2 \cdot \frac{\frac{6}{25} + 2 \cdot \frac{6}{25} + 3 \cdot \frac{4}{25}}{6}} = 0{,}63245;$$

die wahrscheinliche Abweichung, wenn man in dem Ausdrucke für P das zweite Glied unbeachtet läßt, beträgt also

$$\varrho k \sqrt{s} = 0{,}476936 \cdot 0{,}63245 \cdot \sqrt{1200} = 10{,}3 \cdots,$$

so daß die Wiederholungszahl der weißen Kugeln mit der Wahrscheinlichkeit $\frac{1}{2}$ zwischen den Grenzen 430 und 450 zu erwarten ist.

Anders ergeben sich diese Grenzen, wenn die 1200 Ziehungen aus *einer* Urne gemacht werden, welche der weißen Kugel die Wahrscheinlichkeit $\frac{11}{30}$ verleiht, also beispielsweise aus einer Urne mit 11 weißen und 19 schwarzen Kugeln; dann ist nämlich die wahrscheinliche Abweichung

$$0{,}476936\sqrt{2 \cdot 1200 \cdot \frac{11}{30} \cdot \frac{19}{30}} = 11{,}1 \cdots,$$

die wahrscheinlichen Grenzen für die Wiederholungszahl der weißen Kugeln sind also 439 und 451, weiter als im andern Falle.

95. Mittlere Abweichung. Für die Wahrscheinlichkeit einer Abweichung vom absoluten Betrage ξ ergab sich der Näherungsausdruck:

$$\frac{2}{k\sqrt{\pi s}}\, e^{-\frac{\xi^2}{k^2 s}}.$$

Multipliziert man diesen Ausdruck mit ξ^2 und integriert, ξ als stetige Variable auffassend, zwischen den Grenzen 0 und ∞, so ergibt sich das mittlere Quadrat der Abweichung l, nämlich:

$$\begin{aligned} \mathfrak{D}(l^2) &= \frac{2}{k\sqrt{\pi s}} \int_0^\infty \xi^2 e^{-\frac{\xi^2}{k^2 s}} d\xi \\ &= \frac{2k^2 s}{\sqrt{\pi}} \int_0^\infty t^2 e^{-t^2} dt \\ &= \frac{k^2 s}{2}, \end{aligned}$$

weil das Integral (s. Nr. 12) den Wert $\frac{\sqrt{\pi}}{4}$ hat. Mit Rücksicht auf den Wert von k, Gleichung (5), ist also

$$\mathfrak{D}(l^2) = \sum p_\lambda q_\lambda; \tag{15}$$

die Quadratwurzel hieraus ist die mittlere Abweichung von der wahrscheinlichsten Wiederholungszahl.

Bestünde hingegen durch die ganze Dauer der Versuche dieselbe Wahrscheinlichkeit für E und so auch für $\bar{E}$, nämlich die durchschnittliche Wahrscheinlichkeit p, beziehungsweise q, so ergäbe sich nach Nr. 80 als mittleres Quadrat der Abweichung:

$$\mathfrak{D}_1(l^2) = spq.$$

Um diese beiden Werte miteinander vergleichen zu können, bringen wir sie auf folgende Gestalt:

$$\mathfrak{D}(l^2) = p_1(1-p_1) + p_2(1-p_2) + \cdots + p_s(1-p_s)$$
$$= p_1 + p_2 + \cdots + p_s - (p_1^2 + p_2^2 + \cdots + p_s^2),$$
$$\mathfrak{D}_1(l^2) = s\frac{p_1 + p_2 + \cdots + p_s}{s}\left(1 - \frac{p_1 + p_2 + \cdots + p_s}{s}\right)$$
$$= p_1 + p_2 + \cdots + p_s - \frac{(p_1 + p_2 + \cdots p_s)^2}{s};$$

mithin ist

$$\mathfrak{D}_1(l^2) - \mathfrak{D}(l^2) = p_1^2 + p_2^2 + \cdots + p_s^2 - \frac{(p_1 + p_2 + \cdots + p_s)^2}{s}$$
$$= \frac{(1^2 + 1^2 + \cdots + 1^2)(p_1^2 + p_2^2 + \cdots + p_s^2) - (p_1 + p_2 + \cdots + p_s)^2}{s}$$
$$= \frac{(p_1 - p_2)^2 + (p_1 - p_3)^2 + \cdots + (p_2 - p_3)^2 + \cdots}{s} > 0,$$

also

$$\mathfrak{D}(l^2) < \mathfrak{D}_1(l^2).$$

Daraus ergibt sich der Satz:

„Die mittlere Abweichung vom wahrscheinlichsten Resultate ist kleiner, die Präzision also größer, wenn die Ziehungen in der beschriebenen Weise aus verschiedenen Urnen vorgenommen werden, als wenn sie aus einer Urne erfolgten, die dem Erscheinen von E die durchschnittliche Wahrscheinlichkeit verleiht."

Würden z Reihen von je s Versuchen auf die erste Art ausgeführt werden, so wäre für die dabei zum Vorschein kommenden Quotienten

$$\frac{m_1}{s}, \quad \frac{m_2}{s}, \ldots \frac{m_z}{s}, \tag{R}$$

welche die relative Häufigkeit des Erscheinens von E in den einzelnen Reihen darstellen, eine andere wahrscheinlichste Verteilung oder *Dispersion* um den Wert p zu erwarten als für die Quotienten

$$\frac{m_1'}{s}, \quad \frac{m_2'}{s}, \ldots \frac{m_z'}{s}, \tag{R'}$$

welche sich ergäben, wenn die z Ziehungsreihen aus einer und derselben Urne vorgenommen worden wären, welche dem E die durchschnittliche Wahrscheinlichkeit p verleiht. Der Unterschied bestünde darin, daß sich die Reihe (R) enger um den Wert p zusammendrängte als die Reihe (R′). Bezeichnet man die wahrscheinlichste Dispersion der Ergebnisse von Reihen, welche nach Art der Reihe (R′), also

unter beständig gleichbleibenden Umständen ausgeführt werden, als *normale Dispersion*, so gebührt der Verteilung der Ergebnisse der Reihe (R) die Bezeichnung *unternormale Dispersion.*[1])

96. Beispiel XLVIII. *Eine Urne enthält σ Kugeln, davon sind σp weiß, σq schwarz. Man macht s ($< \sigma$) Ziehungen, ohne die gezogenen Kugeln zurückzulegen. Wie groß ist unter der Voraussetzung, daß σ, s und $\sigma - s$ große Zahlen seien, die Wahrscheinlichkeit, daß $sp + l$ weiße und $sq - l$ schwarze Kugeln zum Vorschein kommen?*

Der Fall ist dadurch bemerkenswert, daß sich die Wahrscheinlich keiten für das Ziehen einer weißen, resp. schwarzen Kugel, die vor Beginn der Ziehungen p, q sind, im Laufe derselben auf unbekannte Art ändern.

Der strenge Ausdruck für die verlangte Wahrscheinlichkeit ist:

$$T_l = \frac{s!}{(sp+l)!\,(sq-l)!}$$
$$\cdot \frac{\sigma p(\sigma p - 1)\cdots(\sigma p - sp - l + 1)\cdot \sigma q(\sigma q - 1)\cdots(\sigma q - sq + l + 1)}{\sigma(\sigma - 1)\cdots(\sigma - s + 1)};$$

der erste Faktor gibt die Anzahl der Arten an, auf welche das bezeichnete Ereignis eintreten kann, der zweite gibt die Wahrscheinlichkeit bei einer bestimmten Anordnung der Kugeln, die aber für alle Anordnungen die gleiche ist. Nun läßt sich T_l wie folgt durch lauter Fakultäten ausdrücken:

$$T_l = \frac{s!}{(sp+l)!\,(sq-l)!}\,\frac{(\sigma p)!\,(\sigma q)!\,(\sigma - s)!}{[(\sigma - s)p - l]!\,[(\sigma - s)q + l]!\,\sigma!}.$$

Unter den über σ, s und $\sigma - s$ gemachten Voraussetzungen kann näherungsweise gesetzt werden (vgl. Nr. 75):

$$\frac{s!}{(sp+l)!\,(sq-l)!}\,p^{sp+l}\,q^{sq-l} = \frac{1}{\sqrt{2\pi spq}}\,e^{-\frac{l^2}{2spq}}$$

$$\frac{(\sigma - s)!}{[(\sigma - s)p - l]!\,[(\sigma - s)q + l]!}\,p^{(\sigma - s)p - l}\,q^{(\sigma - s)q + l} = \frac{1}{\sqrt{2\pi(\sigma - s)pq}}\,e^{-\frac{l^2}{2(\sigma - s)pq}};$$

daraus ergibt sich durch Multiplikation:

$$\frac{s!}{(sp+l)!\,(sq-l)!}\,\frac{(\sigma - s)!}{[(\sigma - s)p - l]!\,[(\sigma - s)q + l]!}\,p^{\sigma p}\,q^{\sigma q}$$
$$= \frac{1}{2\pi pq\sqrt{s(\sigma - s)}}\,e^{-\frac{\sigma l^2}{2s(\sigma - s)pq}}.$$

1) Der Begriff der Dispersion und ihre Unterscheidung in normale, unternormale und übernormale rührt von W. Lexis her: Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft (1877), p. 22. — Vgl. auch J. v. Kries, Die Prinzipien der Wahrscheinlichkeitsrechnung (1886), p. 108 ff. — Ein Beispiel übernormaler Dispersion folgt in der nächsten Nummer.

Ferner ist nach Nr. 74

$$\frac{\sigma!}{(\sigma p)!\,(\sigma q)!}\,p^{\sigma p}\,q^{\sigma q} = \frac{1}{\sqrt{2\pi\sigma p q}}\,.$$

Durch Division der vorletzten Gleichung durch die letzte erhält man:

$$T_l = \frac{1}{\sqrt{2\pi s \frac{\sigma - s}{\sigma} p q}}\, e^{-\frac{l^2}{2 s \frac{\sigma - s}{\sigma} p q}}\,.$$

Dies ist aber derselbe Ausdruck, welcher sich ergeben hätte, wenn die Wahrscheinlichkeiten p, q im ganzen Verlaufe der Ziehungen konstant blieben, wenn jedoch statt s Versuchen deren bloß $\frac{\sigma - s}{\sigma} s$ gemacht würden. Je größer σ gegenüber s ist, desto geringer der Einfluß, den das Nichtzurücklegen der Kugeln ausübt; für $\lim \sigma = \infty$ wird $\frac{\sigma - s}{\sigma} = 1$, d. h. bei einer unbegrenzten Menge von Kugeln in der Urne hört jener Einfluß völlig auf.

Würde man z Beobachtungsreihen gleichen Umfanges in der beschriebenen Weise ausführen und die Quotienten

$$\frac{m_1}{s},\ \frac{m_2}{s},\ \ldots\ \frac{m_z}{s}$$

bilden, welche die relative Häufigkeit des Erscheinens einer weißen Kugel in den einzelnen Reihen ausdrücken, so wäre gegenüber dem Falle, wo die Wahrscheinlichkeiten p, q im Verlauf der Versuche konstant bleiben, eine *übernormale Dispersion* dieser Quotienten zu erwarten, d. h. eine Verteilung derselben, bei welcher sie sich weniger dicht um den Wert p zusammendrängen als in dem entgegengestellten Falle. Denn die Präzision einer Beobachtungsreihe der oben beschriebenen Art ist

$$\sqrt{\frac{s\frac{\sigma - s}{\sigma}}{2pq}},$$

die einer Beobachtungsreihe mit konstant bleibenden Wahrscheinlichkeiten

$$\sqrt{\frac{s}{2pq}},$$

die letztere also größer als die erstere.

§ 3. Das Gesetz der großen Zahlen.

97. Elementare Wahrscheinlichkeiten und Durchschnittswahrscheinlichkeiten. Die nähere Vergleichung des Bernoulli-

schen Theorems mit dem Poissonschen gibt Anlaß zu folgenden Bemerkungen.

In dem typischen Falle, auf welchen sich das Bernoullische Theorem bezieht, sind die konstant bleibenden Wahrscheinlichkeiten p, q der Ereignisse E, $\bar{E}$ maßgebend sowohl für die Bestimmung der wahrscheinlichsten Kombination von E, $\bar{E}$ in s Versuchen, wie auch für die Präzision der Versuchsreihe und daher auch für die Dispersion der Ergebnisse wiederholter derartiger Versuchsreihen gleichen Umfanges.

Anders verhält es sich in dem Falle, welcher dem Poissonschen Theorem zugrunde liegt. Hier gibt es wohl zwei Größen p, q, nämlich die arithmetischen Mittel der p_λ und q_λ, die für die wahrscheinlichste Kombination von E und $\bar{E}$ bestimmend sind in derselben Weise wie in dem ersten Falle, sie sind aber nicht maßgebend für die Beurteilung der Präzision und Dispersion; diese hängen vielmehr von den einzelnen p_λ, q_λ ab und können sich daher bei demselben p und q sehr verschieden gestalten.

Wir wollen diesen Unterschied vorläufig dadurch zum Ausdruck bringen, daß wir p, q im ersten Falle als *Elementarwahrscheinlichkeiten*, im zweiten Falle als *Durchschnittswahrscheinlichkeiten* bezeichnen. Das verschiedene Verhalten beider veranlaßt uns, dem Wesen der Durchschnittswahrscheinlichkeiten näher zu treten. Die folgende Betrachtung dürfte einen hierzu geeigneten Ausgangspunkt darbieten.

Eine Urne U enthalte N Kugeln, wovon G weiß, die übrigen schwarz sein mögen; die Wahrscheinlichkeit für das Ziehen einer weißen Kugel (Ereignis E) ist

$$p = \frac{G}{N}.$$

Nun stelle man sich vor, daß die Kugeln durch irgendein Merkmal[1]) in ν Kategorien unterschieden seien, welche der Reihe nach aus $N_1, N_2, \cdots N_\nu$ Kugeln bestehen, worunter sich $G_1, G_2, \cdots G_\nu$ weiße befinden mögen; dann ist $N_1 + N_2 + \cdots + N_\nu = N$ und $G_1 + G_2 + \cdots + G_\nu = G$, und die Wahrscheinlichkeit von E läßt sich in der Form

$$p = \frac{G_1 + G_2 + \cdots + G_\nu}{N} = \frac{N_1}{N}\frac{G_1}{N_1} + \frac{N_2}{N}\frac{G_2}{N_2} + \cdots + \frac{N_\nu}{N}\frac{G_\nu}{N_\nu}$$

darstellen. Dieser Ausdruck gestattet folgende Deutung: Man kann jetzt von ν Modalitäten des Ereignisses E (und seines Gegensatzes $\bar{E}$) reden; $\frac{N_1}{N} = \omega_1$ ist die Wahrscheinlichkeit, daß die erste Modalität sich

1) Z. B. durch aufgeschriebene Nummern.

einstelle; $\frac{G_1}{N_1} = p_1$ die Wahrscheinlichkeit, daß dann das Ereignis E eintreffe[1]). Mit diesen Bezeichnungen schreibt sich p entweder

$$p = \omega_1 p_1 + \omega_2 p_2 + \cdots + \omega_\nu p_\nu$$

oder

$$p = \frac{N_1 p_1 + N_2 p_2 + \cdots + N_\nu p_\nu}{N},$$

und erscheint bei der letzteren Schreibweise als arithmetisches Mittel der den einzelnen Modalitäten entsprechenden Wahrscheinlichkeiten, jede Wahrscheinlichkeit mit dem der Modalität entsprechenden Gewichte genommen, also in der Form einer Durchschnittswahrscheinlichkeit.

Man kann, um das Verhältnis der $p_1, p_2, \cdots p_\nu$ zu dem p zu kennzeichnen, die ersteren *Spezial-* oder *Partial*wahrscheinlichkeiten, p dagegen eine *General-* oder *Total*wahrscheinlichkeit nennen.[2])

Wie auch die Zerlegung in Kategorien erfolgen möge, an dem Wesen der Frage nach der Wahrscheinlichkeit des Erscheinens einer weißen Kugel wird nichts geändert, und gerade hierin ist das charakteristische Merkmal einer elementaren Wahrscheinlichkeit zu erblicken.

Nun aber denke man sich die Kugelkategorien voneinander getrennt, indem man jede in eine besondere Urne legt; an die Stelle der einen Urne U treten dann ν Urnen $U_1, U_2, \cdots U_\nu$, welche dem Ereignis E der Reihe nach die Wahrscheinlichkeit $p_1, p_2, \cdots p_\nu$ verleihen.

Die Frage nach der Wahrscheinlichkeit des Ereignisses E kann jetzt verschiedene Beantwortung erfahren je nach den Bestimmungen, welche über die Ausführung der Ziehungen getroffen werden.

a) Setzt man fest, daß aus einer *bestimmten* Urne, z. B. aus U_1, gezogen werden soll, so hat E die Wahrscheinlichkeit p_1.

b) Bestimmt man, daß einer *beliebigen* Urne eine Kugel entnommen werden soll, so ist

$$p' = \frac{1}{\nu} p_1 + \frac{1}{\nu} p_2 + \cdots + \frac{1}{\nu} p_\nu = \frac{p_1 + p_2 + \cdots + p_\nu}{\nu}$$

die Wahrscheinlichkeit, daß sie weiß sein werde.

c) Man kann die Bestimmungen auch so treffen, daß bezüglich der Erwartungsbildung wieder die ursprünglichen Verhältnisse ein-

1) Poisson, Recherches etc., deutsche Übersetz., p. 100 und 115, bezeichnet die Modalitäten als „Ursachen" der Ereignisse E (oder $\bar{E}$) und belegt die Zahlenreihe $\omega_1, \omega_2, \cdots \omega_\nu$ mit dem Namen „Wahrscheinlichkeitsgesetz der Ursachen".

2) J. v. Kries, Prinz. d. Wahrsch.-R., p. 106. — L. v. Bortkewitsch, Krit. Betracht. z. theoret. Statist., Jahrb. f. Nationalök. u. Stat. (3) 8 (1894), p. 642.

treten. Man lege in eine Hilfsurne $N_1 \lambda$ Zettel mit der Nummer 1, $N_2 \lambda$ Zettel mit der Nummer 2, $\cdots$, $N_\nu \lambda$ Zettel mit der Nummer ν (λ bedeutet eine beliebige ganze Zahl) und ziehe zuerst aus dieser Urne einen Zettel; die Nummer desselben bezeichnet die Urne, aus welcher die Kugel zu ziehen ist. Unter diesen Umständen ist die Wahrscheinlichkeit für das Erscheinen einer weißen Kugel wieder

$$p = \frac{N_1}{N}\frac{G_1}{N_1} + \frac{N_2}{N}\frac{G_2}{N_2} + \cdots + \frac{N_\nu}{N}\frac{G_\nu}{N_\nu}.$$

Auch dadurch würde dieses Ziel erreicht, daß man, statt einer, $N_1 \lambda$ Urnen von der Sorte U_1, $N_2 \lambda$ Urnen von der Sorte U_2, $\cdots$, $N_\nu \lambda$ Urnen von der Sorte U_ν, im ganzen $N\lambda$ Urnen aufstellt und aus einer beliebigen die Kugel zieht.

Nun handle es sich nicht um die einzelne Ziehung, sondern um eine Reihe, sagen wir von $s = N\lambda$ Ziehungen. Wir wollen zwei Ausführungsmodalitäten einander gegenüberstellen.

α) Man benütze das System U_1, U_2, $\cdots$ U_ν von ν Urnen, mache $N_1 \lambda$ Ziehungen aus U_1, $N_2 \lambda$ Ziehungen aus U_2, $\cdots$, schließlich $N_\nu \lambda$ Ziehungen aus U_ν, die gezogene Kugel jedesmal zurücklegend und unter die andern mischend. Die Durchschnittswahrscheinlichkeit, welche bei diesem Vorgange der weißen Kugel entspricht, ist

$$p = \frac{N_1 p_1 + N_2 p_2 + \cdots + N_\nu p_\nu}{N}.$$

β) Man benütze das unter c) erwähnte System von $N\lambda$ Urnen, worunter sich $N_1 \lambda$ von der Sorte U_1 befinden usw., und mache jede der s Ziehungen aus einer beliebigen Urne des Systems, die Kugel jedesmal zurücklegend; dabei wird zugleich vorausgesetzt, daß die Urnen beständig durcheinander gemengt werden, daß man also nicht wisse, aus welchen bereits gezogen worden ist. Auch bei diesem Vorgange ist die Durchschnittswahrscheinlichkeit für eine weiße Kugel

$$p = \frac{N_1 p_1 + N_2 p_2 + \cdots + N_\nu p_\nu}{N}.$$

Zwischen diesen an Wert gleichen Durchschnittswahrscheinlichkeiten besteht aber ein fundamentaler Unterschied.

Während man im Falle α) mit Sicherheit weiß, daß genau N_1-mal die Spezialwahrscheinlichkeit p_1, genau $N_2 \lambda$-mal die Spezialwahrscheinlichkeit p_2, usf. zur Geltung kam, läßt sich im Falle β) keine bestimmte Aussage hierüber machen. Zwar ist, sofern s eine sehr große Zahl vorstellt, dem Bernoullischen Theorem zufolge mit großer Wahrscheinlichkeit zu erwarten, daß die relative Häufigkeit, mit welcher die Urnen der einzelnen Gattungen daran kommen, nicht all-

zuviel von den Verhältnissen $\frac{N_1}{N}, \frac{N_2}{N}, \cdots$ abweichen werde; es liegen aber auch erhebliche Abweichungen von diesen Verhältnissen im Bereiche der Möglichkeit.

Im Falle α) kann von einer Wahrscheinlichkeit des Ereignisses E schlechtweg, d. i. für die ganze Dauer der Versuche, nicht gesprochen werden; sie ist eben $N_1\lambda$-mal p_1, $N_2\lambda$-mal p_2 usf. Im Falle β) aber gibt p die Wahrscheinlichkeit von E für die ganze Versuchsdauer an; damit steht keineswegs im Widerspruche, daß bei der Ausführung des einzelnen Versuches, nachdem man sich bereits für die Urne entschieden hat, aus welcher der nächste Zug geschehen soll, eine der Spezialwahrscheinlichkeiten wirksam wird; bei der völligen Unwissenheit über die Anordnung der Urnen besteht doch immer nur die Generalwahrscheinlichkeit dafür, daß man eine weiße Kugel ziehen werde.

Der Fall β) ist also völlig äquivalent der Ausführung von s Ziehungen aus der einen ursprünglichen Urne U, welche dem Erscheinen einer weißen Kugel die Elementarwahrscheinlichkeit p gab. Der Fall α) aber läßt sich in keiner Weise durch Ziehungen aus einer einzigen Urne von gleichbleibender Füllung ersetzen.

Man hat, um das Ergebnis der Betrachtung zusammenzufassen, zwei Arten von Durchschnittswahrscheinlichkeiten zu unterscheiden: solche von *bestimmter* und solche von *willkürlicher* Zusammensetzung, oder in anderer Terminologie[1]): konstant zusammengesetzte Durchschnittswahrscheinlichkeiten und Durchschnittswahrscheinlichkeiten im eigentlichen Sinne. Die letzteren sind analog den elementaren Wahrscheinlichkeiten.

Daraus ergibt sich als naturgemäße Folgerung, daß auf Versuchsreihen, welchen eine elementare *oder* eine Durchschnittswahrscheinlichkeit im eigentlichen Sinne zugrunde liegt, das Bernoullische Theorem zur Anwendung kommen kann, daß also, was den letzteren Fall betrifft, die Präzision lediglich nach der Durchschnittswahrscheinlichkeit zu beurteilen ist. Bei Versuchsreihen hingegen, welchen eine Durchschnittswahrscheinlichkeit von bestimmter Zusammensetzung zugrunde liegt, kommt das Theorem von Poisson zur Geltung; hier richtet sich die Präzision nach der Zusammensetzung der Durchschnittswahrscheinlichkeit.

98. Das Gesetz der großen Zahlen. Der theoretische Satz über den Zusammenhang zwischen der relativen Häufigkeit des Eintreffens eines Ereignisses in einer großen Anzahl von Versuchen und seiner Wahrscheinlichkeit, der in Nr. 87 unter der einschränkenden Voraussetzung formuliert worden ist, daß diese Wahrscheinlichkeit

---

1) L. v. Bortkewitsch, l. c., p. 650—651.

von Versuch zu Versuch dieselbe bleibe, läßt sich jetzt auch auf solche Fälle ausdehnen, wo dem Ereignis eine Durchschnittswahrscheinlichkeit, eine konstant zusammengesetzte oder eine eigentliche, zukommt. Man kann nämlich mit Rücksicht auf die letzten Ergebnisse sagen:

„Wenn zwei einander ausschließenden Ereignissen E, $\bar{E}$ solche im Laufe der Verwirklichungen unveränderlich bleibende Umstände zugrunde liegen, daß sich für E eine Elementar- oder eine Durchschnittswahrscheinlichkeit p bestimmen läßt, so ist man imstande die Wahrscheinlichkeit P anzugeben, mit welcher erwartet werden darf, daß der Unterschied zwischen der relativen Häufigkeit von E in s Versuchen, $H(E, s)$, und seiner Wahrscheinlichkeit p dem Betrage nach über eine bestimmte Grenze nicht hinausfalle, und wie nahe auch P der Einheit gewählt wird, immer ist

$$\lim_{s=\infty} H(E, s) = p.\text{“}$$

Was die ziffermäßige Ermittlung der Grenze und des ihr entsprechenden P anlangt, so ist in der vorigen Nummer auseinander gesetzt, wann hierzu die Formeln des Bernoullischen und wann jene des Poissonschen Theorems anzuwenden sind.

Die durch vielfältige Erfahrung auch in der Wirklichkeit bestätigte Geltung obigen Satzes in dem Sinne, daß mit wachsendem s die Differenz $H(E, s) - p$ *im allgemeinen* der Null sich nähert, wird als das *Gesetz der großen Zahlen* bezeichnet.

Poisson, der Urheber dieser Bezeichnung[1]), wollte sie nur auf solche Ereignisse angewendet wissen, deren Wahrscheinlichkeit sich im Laufe der Versuche oder Beobachtungen ändert, und er vermeinte, in seinem Theorem, das die theoretische Basis dieses Gesetzes ausmacht, und in dem er eine unmittelbare Aussage über das wirkliche Geschehen erblickte[2]), eine weittragende Verallgemeinerung des Bernoullischen Theorems gefunden zu haben. Wenn man jedoch seine mathematische Deduktion und die zu ihr beigebrachten Illustrationsbeispiele nachprüft, so kommt man zu der Erkenntnis, daß es sich um Ereignisse mit einer Durchschnittswahrscheinlichkeit im eigentlichen Sinne handelt, und auf solche kann ebensogut das Bernoullische Theorem wie das Poissonsche zur Anwendung gebracht werden.[3])

---

1) Comptes rendus 1 (1835), p. 478 sq.

2) Comptes rendus 2 (1836), p. 377 sq.

3) Vgl. die eingehende Beleuchtung dieses Gegenstandes bei L. v. Bortkewitsch, l. c., p. 653 ff.

III. Abschnitt. Wahrscheinlichkeiten auf Grund der Erfahrung.

§ 1. Wahrscheinlichkeit der möglichen Ursachen eines beobachteten Ereignisses.

99. Entwicklung der Fragestellung. Wenn über zwei Ereignisse (Tatbestände) E und $\bar{E}$, die einander gegenseitig ausschließen, ein solches Wissen zu Gebote steht, das gestattet, die Umfänge der Realisierungsmöglichkeiten, die mit Notwendigkeit das eine oder das andere Ereignis herbeiführen, gegeneinander abzugrenzen und miteinander quantitativ zu vergleichen, so lassen sich auf Grund dieses Wissens allein die für die Erwartungsbildung maßgebenden Wahrscheinlichkeiten berechnen, sei es, daß es sich nur um eine einzelne Realisierung oder um eine Folge von solchen handelt.

Die hierdurch gekennzeichnete allgemeine Methode wird als *Wahrscheinlichkeitsbestimmung a priori*, eine nach dieser Methode bestimmte Wahrscheinlichkeit auch als apriorische Wahrscheinlichkeit bezeichnet. Bisher waren es nur Wahrscheinlichkeiten dieser Art, die uns beschäftigt haben.

Es bedeutet nun einen wichtigen Schritt in der Entwicklung der Wahrscheinlichkeitstheorie, daß man sie auch auf Ereignisse auszudehnen sucht, bezüglich welcher ein so geartetes Wissen nicht vorliegt, wo jedoch der beobachtete Erfolg einer oder mehrerer Realisierungen zu Gebote steht. Denn die Anwendungen der Wahrscheinlichkeitstheorie auf Naturerscheinungen und Vorgänge des praktischen Lebens bieten fast ausschließlich Fälle dieser Art dar.

Bevor es hier zu Erwartungsbildungen in bezug auf künftige Erfolge kommen kann, muß ein Rückschluß von dem beobachteten Erfolge auf die ihn bedingenden Umstände gemacht werden, und dieser beruht wieder auf einem Wahrscheinlichkeitsurteil. Wir nehmen an, das vorhandene, aber unzureichende Wissen in Verbindung mit dem beobachteten Ereignis gestatte die Aufstellung einer begrenzten oder auch unbegrenzten Menge von Annahmen über die dem beobachteten Ereignis zugrunde liegenden Umstände oder Bedingungen. Sowie nun ein bestimmter, bezeichneter Erfolg aus verschiedenen Bedingungskomplexen mit verschiedenem Grade der Wahrscheinlichkeit zu erwarten ist, so kommt auch den verschiedenen möglichen Bedingungskomplexen, aus welchen ein wirklich beobachteter Erfolg hervorgegangen sein kann, verschiedene Wahrscheinlichkeit der Existenz zu. Erst wenn die Messung dieser gelungen ist, kann eine Berechnung über die Wahrscheinlichkeit erst zu erwartender Erfolge angestellt werden.

Die Methode, welche ein zur apriorischen Wahrscheinlichkeitsbestimmung unzureichendes Wissen und die Ergebnisse der Erfahrung,

der Beobachtung, zur Lösung von Aufgaben verwendet, wird als *Wahrscheinlichkeitsbestimmung a posteriori* bezeichnet. Unter einer aposteriorischen Wahrscheinlichkeit ist daher eine solche zu verstehen, bei deren Berechnung auch die Erfahrung mitgewirkt hat.

Es ist in der Wahrscheinlichkeitstheorie üblich geworden, die für die Erwartungsbildung maßgebenden Umstände im Hinblick auf ein ungewisses Ereignis seine *Ursache* zu nennen. In diesem Sinne spricht man auch von den möglichen Ursachen eines beobachteten Ereignisses und meint darunter die verschiedenen Bedingungskomplexe, aus welchen das Ereignis hervorgegangen sein kann. Es besteht hier ein Widerspruch mit dem gewöhnlichen Sprachgebrauch, indem unter Ursache nicht das verstanden wird, was dem Ereignis das Eintreffen, die Existenz, mit Notwendigkeit verleiht, sondern etwas, was sie ihm verleihen *kann*. Am zutreffendsten wäre es wohl, von Annahmen oder *Hypothesen* über die dem beobachteten Ereignis zugrunde liegenden Umstände zu sprechen; in vielen Fällen erweist sich die Bezeichnung „Entstehungsmodi" des beobachteten Ereignisses als zutreffend.[1])

Wir gehen nun daran, dasjenige Theorem in seinen verschiedenen Formen abzuleiten, auf welchem die Beurteilung der Wahrscheinlichkeit der verschiedenen Ursachen beruht, denen ein beobachtetes Ereignis zugeschrieben werden kann. Die Ableitung soll an schematischen Problemen vorgenommen werden, um deutlich erkennen zu lassen, daß sie keiner neuen Prinzipe bedarf und lediglich auf der Wahrscheinlichkeitsdefinition beruht, mit andern Worten, daß die aposteriorische Wahrscheinlichkeitsrechnung mit denselben logischen Schlüssen operiert wie die apriorische und nur in den Grundlagen der Rechnung von dieser sich unterscheidet. Einer späteren Stelle bleibt es dann vorbehalten, die Voraussetzungen, welche gemacht werden, auf ihren wesentlichen Inhalt zu prüfen und zu untersuchen, wieweit sie in besonderen Fällen erfüllt sind.

100. Theorem von Bayes[2]) für den Fall, daß die Ursachen a priori gleichmöglich sind.

Wir stellen folgendes Problem zur Lösung:

„Es liegen n äußerlich gleiche Urnen U_1, U_2, ... U_n von folgender Füllung vor:

---

1) Vgl. hierzu J. v. Kries, Prinz. d. Wahrscheinlichkeitsrechnung, p. 122, und K. Stumpf, Über den Begriff d. mathem Wahrsch. Sitzungsber. d. Münch. Ak. 1892, p. 96ff.

2) J. Bayes war der erste, der sich (in einer nach seinem Ableben durch Price veröffentlichten Abhandlung, erschienen im 53. Bande der Philos. Trans. [1764], p. 370—418) mit der Beurteilung der Wahrscheinlichkeit von Ursachen beschäftigte. Die Aufstellung und Begründung der diesen Teil der Theorie beherrschenden Sätze verdankt man Laplace (Théorie analyt).

U_1 enhält c_1 Kugeln, darunter a_1 weiße;
U_2 „ c_2 „ „ a_2 „ ;
. . . .
U_n „ c_n „ „ a_n „ ;

man hat aus einer der Urnen eine Kugel gezogen und sie weiß gefunden. Wie groß ist die Wahrscheinlichkeit, daß die Kugel aus der Urne U_i stamme?“

Für das beobachtete Ereignis gibt es n mögliche Entstehungsmodi oder Ursachen, dargestellt durch die n Urnen; die Ursachen sind a priori, d. h. vor der Ziehung, *gleichmöglich*, weil für jede Urne die Wahrscheinlichkeit $\frac{1}{n}$ besteht, daß man sie wählen werde.

Um die Lösung der Aufgabe auf die Bestimmung einer Wahrscheinlichkeit im bisher geläufigen Sinne zurückzuführen, gestalten wir die Bedingungen folgendermaßen um.

Zunächst bringen wir die Urnen auf gleiche Kugelzahl, ohne das Mischungs*verhältnis* in Ansehung der weißen Kugeln zu ändern. Ist v das kleinste gemeinschaftliche Vielfache (oder *ein* Vielfaches) der Zahlen $c_1, c_2, \ldots c_n$, und ist

$$v = \gamma_1 c_1 = \gamma_2 c_2 = \cdots = \gamma_n c_n, \tag{1}$$

so enthalten die Urnen je v Kugeln und darunter

$$\gamma_1 a_1, \quad \gamma_2 a_2, \quad \cdots \quad \gamma_n a_n$$

weiße.

Nun dürfen, ohne daß die ursprüngliche Gleichmöglichkeit der Einzelfälle gestört würde, die Inhalte der Urnen in *einer* Urne U vereinigt werden; vorher jedoch sollen die Kugeln jeder Urne mit der gleichen Nummer versehen werden, welche die Urne trägt, um sie auch jetzt noch nach ihrer Abstammung unterscheiden zu können.

Die ursprünglich gestellte Frage ist jetzt identisch mit der folgenden: Eine aus der Urne U gezogene Kugel war weiß; wie groß ist die Wahrscheinlichkeit, daß sie mit der Nummer i bezeichnet sei?

Aber auch diese Frage kann vermöge des Umstandes, daß die Farbe der gezogenen Kugel bekannt ist, durch eine andere ersetzt werden: Wie groß ist die Wahrscheinlichkeit, aus einer Urne U', in die man aus U nur die weißen Kugeln gelegt hat, eine Kugel mit der Nummer i zu ziehen?

Und die Antwort hierauf ist auf Grund der Wahrscheinlichkeitsdefinition gegeben durch:

$$P_i = \frac{\gamma_i a_i}{\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_n a_n}.$$

Ersetzt man die γ durch ihre Werte aus der Relation (1), so wird

$$P_i = \frac{v\frac{a_i}{c_i}}{v\frac{a_1}{c_1} + v\frac{a_2}{c_2} + \cdots + v\frac{a_n}{c_n}} = \frac{\frac{a_i}{c_i}}{\frac{a_1}{c_1} + \frac{a_2}{c_2} + \cdots + \frac{a_n}{c_n}};$$

nun ist aber $\frac{a_i}{c_i} = p_i$ die Wahrscheinlichkeit, aus der Urne U_i eine weiße Kugel zu ziehen; demnach ist schließlich die Lösung für das ursprünglich gestellte Problem:

$$P_i = \frac{p_i}{p_1 + p_2 + \cdots + p_n}. \tag{2}$$

Abstrahiert man von dem zugrunde gelegten Urnenschema, so ergibt sich der einfachste Ausdruck des Bayesschen *Prinzips* in dem folgenden Satze:

„Wenn ein beobachtetes Ereignis E mehreren a priori gleichmöglichen Ursachen zugeschrieben werden kann, jedoch so, daß eine von ihnen notwendig wirksam war, so ist die Wahrscheinlichkeit, daß die Ursache U_i es war, gleich der Wahrscheinlichkeit, mit welcher das Ereignis aus ihr zu erwarten ist, geteilt durch die Summe der auf alle Ursachen bezogenen Wahrscheinlichkeiten dieses Ereignisses."

Eine besonders durchsichtige Darstellung erhält der Satz durch Heranziehung des Begriffs der relativen Wahrscheinlichkeit; denn alle in (2) vorkommenden Buchstaben bedeuten relative Wahrscheinlichkeit, u. zw.

p_i die Wahrscheinlichkeit von E unter der Voraussetzung, daß U_i wirksam ist, also $\mathfrak{W}_{U_i}(E)$,

P_i die Wahrscheinlichkeit von U_i unter der Voraussetzung, daß E eingetreten ist, also $\mathfrak{W}_E(U_i)$;

infolgedessen kann (2) in der Gestalt geschrieben werden:

$$\mathfrak{W}_E(U_i) = \frac{\mathfrak{W}_{U_i}(E)}{\sum\limits_{i}^{n} \mathfrak{W}_{U_i}(E)}. \tag{2*}$$

Hieraus ergibt sich die Bemerkung, daß die Wahrscheinlichkeit einer Ursache proportional ist der Wahrscheinlichkeit, welche sie dem beobachteten Ereignis verleiht; daraus folgt weiter, daß die größte Wahrscheinlichkeit jener Ursache zukommt, aus welcher unter allen das beobachtete Ereignis mit der größten Wahrscheinlichkeit zu erwarten ist. Die Grenze dieses Zusammenhanges ist der *Kausalnexus*: Wo Ereignis und Ursache durch das Prädikat der (gegenseitigen) Notwendigkeit miteinander verknüpft sind, ist aus der Existenz des

Ereignisses auch auf die Existenz der Ursache mit Sicherheit zu schließen.

101. Theorem von Bayes für den Fall, daß die Ursachen a priori verschiedene Wahrscheinlichkeit besitzen.

Es sei das folgende Problem zu lösen: „Eine Urne enthält c Kugeln; davon sind c_1 mit der Nummer 1 bezeichnet und a_1 davon weiß; c_2 mit der Nummer 2 bezeichnet und a_2 davon weiß usf. bis zur Nummer n. Es ist eine Ziehung gemacht und konstatiert worden, daß die gezogene Kugel weiß ist. Wie groß ist die Wahrscheinlichkeit, daß sie die Nummer i trage?"

Für das beobachtete Ereignis gibt es n Entstehungsmodi, dargestellt durch die Gruppen gleichbezeichneter Kugeln; weil diese Gruppen ungleich zahlreich sind, so bestehen a priori verschiedene Wahrscheinlichkeitsgrade der Ursachen.

Mit Rücksicht darauf, daß die Farbe der gezogenen Kugel bekannt ist, kann die Aufgabe in eine andere umgewandelt werden, indem man die weißen Kugeln, in der Anzahl

$$a = a_1 + a_2 \cdots + a_n$$

vorhanden, in eine andere Urne legt und nun die Frage nach der Wahrscheinlichkeit stellt, daß ein Zug eine mit der Nummer i bezeichnete Kugel bringe.

Diese Wahrscheinlichkeit ist aber

$$P_i = \frac{a_i}{a} = \frac{a_i}{a_1 + a_2 + \cdots + a_n};$$

ihr Ausdruck läßt sich wie folgt umformen:

$$P_i = \frac{\frac{c_i}{c}\frac{a_i}{c_i}}{\frac{c_1}{c}\frac{a_1}{c_1} + \frac{c_2}{c}\frac{a_2}{c_2} + \cdots + \frac{c_n}{c}\frac{a_n}{c_n}};$$

nun ist $\frac{c_i}{c} = \omega_i$ die Wahrscheinlichkeit, eine Kugel aus der mit der Nummer i versehenen Gruppe zu ergreifen, also die apriorische Wahrscheinlichkeit der Ursache, um deren Wahrscheinlichkeit a posteriori gefragt wird, und $\frac{a_i}{c_i} = p_i$ die Wahrscheinlichkeit, welche diese Ursache dem beobachteten Ereignis verleiht. Die Lösung des gestellten Problems ist also durch

$$P_i = \frac{\omega_i p_i}{\omega_1 p_1 + \omega_2 p_2 + \cdots + \omega_n p_n} \tag{3}$$

gegeben.

Das in der vorigen Nummer formulierte Theorem erfährt also eine Modifikation dahin, daß die Wahrscheinlichkeit des beobachteten Ereignisses nach einer Ursache jedesmal mit der apriorischen Wahrscheinlichkeit der Ursache multipliziert in die Formel eingeht.[1])

Man kann die Formel (3) auch so begründen. Es handelt sich um das Zusammentreffen des beobachteten Ereignisses und derjenigen Ursache, deren Wahrscheinlichkeitsgrad man bestimmen will. Faßt man dieses zusammengesetzte Ereignis zuerst so auf, daß die Ursache existent wird und durch sie das beobachtete Ereignis zustande kommt, so schreibt sich seine Wahrscheinlichkeit:

$$\omega_i p_i;$$

denn vor dem Versuch ist ω_i die Wahrscheinlichkeit der Ursache, welche dem beobachteten Ereignis die Wahrscheinlichkeit p_i erteilt. Stellt man sich dann auf den Standpunkt, daß, wenn das beobachtete Ereignis eintrat, dies infolge der Wirkung der betreffenden Ursache geschah, so ergibt sich für dasselbe zusammengesetzte Ereignis die Wahrscheinlichkeit:

$$(\omega_1 p_1 + \omega_2 p_2 + \cdots + \omega_n p_n) P_i;$$

denn $\omega_1 p_1 + \omega_2 p_2 + \cdots + \omega_n p_n$ ist die Wahrscheinlichkeit, welche dem beobachteten Ereignis überhaupt, ohne Kenntnis der aktiven Ursache, eigen ist, und P_i die Wahrscheinlichkeit, daß dieses Ereignis, nachdem es eingetreten, der bezeichneten Ursache zuzuschreiben sei. Aus der Gleichsetzung der beiden Ausdrücke ergibt sich wieder die Formel (3).

Zu den relativen Wahrscheinlichkeiten in der Formel (2*) treten in dem gegenwärtigen Falle noch die absoluten Wahrscheinlichkeiten $\mathfrak{W}(U_i)$ der Ursachen vor Bekanntwerden des beobachteten Ereignisses, so daß die Formel (3) auch in der Gestalt

$$\mathfrak{W}_E(U_i) = \frac{\mathfrak{W}(U_i)\,\mathfrak{W}_{U_i}(E)}{\sum_1^n{}^i\,\mathfrak{W}(U_i)\,\mathfrak{W}_{U_i}(E)} \qquad (3^*)$$

geschrieben werden kann, aus der ihr logischer Inhalt unmittelbar hervortritt.

---

1) K. Stumpf, l. c., p. 96, hat für die verschiedenen hier auftretenden Wahrscheinlichkeiten eine Nomenklatur vorgeschlagen: Er nennt ω_i die vorgängige Wahrscheinlickkeit, p_i den Erklärungswert, das Produkt $\omega_i p_i$ die abstrakte (i. e. von den übrigen Ursachen absehende) und P_i die konkrete Wahrscheinlichkeit der betreffenden Ursache.

102. Die verschiedenartige Natur der Ursachen. In den Problemen, auf welche die Ableitung der Bayesschen Regel gegründet worden ist, waren die möglichen Ursachen materiell existent, das eine Mal vertreten durch die verschieden gefüllten Urnen, das andere Mal durch die verschieden bezifferten Kugelgruppen in einer Urne. Wahrscheinlichkeit einer Ursache bedeutete hier nicht die Wahrscheinlichkeit, daß sie existent sei, sondern daß das beobachtete Ereignis aus ihr hervorgegangen war.

Anders liegen die Dinge in dem folgenden Problem, das für viele Fälle der Anwendung typisch ist: „Eine Urne enthält c Kugeln, die nur weiß oder schwarz sein können; über den Hergang der Füllung der Urne ist nichts bekannt. Dagegen weiß man, daß in s Ziehungen, wobei die gezogene Kugel jedesmal zurückgelegt und unter die andern gemengt worden ist, immer eine weiße Kugel erschienen ist. Wie groß ist die Wahrscheinlichkeit, daß die Urne a) nur weiße, b) mehr weiße als schwarze Kugeln enthalte?"

Hier sind es die verschiedenen *Annahmen*, welche über das Verhältnis der weißen und schwarzen Kugeln in der Urne gemacht werden können, die man als die möglichen Ursachen des beobachteten Ereignisses bezeichnet.

Der hypothetische Charakter der Ursachen tritt hier deutlich zutage: nur *eine* der Ursachen, d. h. nur *ein* bestimmtes Füllungsverhältnis kann existent sein; die Wahrscheinlichkeit einer Ursache kann hier als Wahrscheinlichkeit ihrer Existenz ausgelegt werden.

In den Beispielen der Nrn. 100 und 101 konnte die Ursache als ein Ereignis, als ein Geschehen aufgefaßt werden: Greifen in die betreffende Urne, Treffen der mit einer bestimmten Nummer versehenen Kugelgruppe.

In dem gegenwärtigen Falle haben die Ursachen nicht die Bedeutung von Ereignissen, sondern die von hypothetischen Tatbeständen.

Wenn hier aus dem vollständigen Nichtwissen über die Entstehungsweise der Urne die Gleichmöglichkeit aller Annahmen über das Füllungsverhältnis deduziert wird, so ruht die gleiche Wahrscheinlichkeit aller Ursachen auf einer andern Grundlage als in dem Beispiel der Nr. 100: dort war sie durch die Existenz und die äußerliche Gleichheit der verschieden gefüllten Urnen objektiv begründet; hier hat es nur eine logische Berechtigung, wenn aus dem völligen Nichtwissen nach dem Prinzip des mangelnden Grundes auf die gleiche Zulässigkeit der möglichen Füllungsverhältnisse geschlossen wird.

Bei der Lösung selbst ist zu beachten, daß mit Rücksicht auf die vorliegende Erfahrung c Annahmen über die Füllung der Urne gemacht werden können, die enthalten sind in dem Schema:

$$\lambda \text{ weiße}, \qquad c - \lambda \text{ schwarze Kugeln} \qquad (\lambda = 1, 2, \cdots c);$$

diesen Annahmen entsprechen die Wahrscheinlichkeiten

$$p_\lambda = \left(\frac{\lambda}{c}\right)^s \qquad (\lambda = 1, 2, \cdots c)$$

des beobachteten Ereignisses.

Die Frage a) ist nach der Wahrscheinlichkeit der Existenz der letzten Füllung gerichtet und hat

$$P_c = \frac{\left(\frac{c}{c}\right)^s}{\left(\frac{1}{c}\right)^s + \left(\frac{2}{c}\right)^s + \cdots + \left(\frac{c}{c}\right)^s} = \frac{c^s}{1^s + 2^s + \cdots + c^s}$$

zur Lösung.

Die Frage b) ist nach einer vollständigen Wahrscheinlichkeit gerichtet und daher durch die Summe der P_λ beantwortet, in welchen λ größer als $c - \lambda$ ist; daher hat man:

für ein gerades c:

$$P = \frac{\left(\frac{c}{2}+1\right)^s + \left(\frac{c}{2}+2\right)^s + \cdots + c^s}{1^s + 2^s + \cdots + c^s},$$

für ein ungerades c:

$$P = \frac{\left(\frac{c+1}{2}\right)^s + \left(\frac{c+3}{2}\right)^s + \cdots + c^s}{1^s + 2^s + \cdots + c^s}.$$

So wäre beispielsweise für $c = 5$ und $s = 4$:

$$P_5 = \frac{5^4}{1^4 + 2^4 + 3^4 + 4^4 + 5^4} = \frac{625}{979} = 0{,}6384,$$

$$P = \frac{3^4 + 4^4 + 5^4}{1^4 + 2^4 + 3^4 + 4^4 + 5^4} = \frac{962}{979} = 0{,}9826.$$

Während nach der obigen Fassung von dem Inhalte der Urne außer der Gesamtzahl der Kugeln nichts bekannt war, nehmen wir jetzt an, derselbe sei so entstanden, daß man aus einer Hilfsurne mit α weißen und β schwarzen Kugeln c-mal zog (die Kugel zurücklegend) und in die zu bildende Urne jedesmal eine Kugel von der Farbe der gezogenen einlegte.

Die c Füllungsmodi haben nun a priori verschiedene Wahrscheinlichkeit; allgemein kommt der Annahme, daß λ weiße und $c - \lambda$ schwarze Kugeln vorhanden sind, die Wahrscheinlichkeit

$$\omega_\lambda = \binom{c}{\lambda}\left(\frac{\alpha}{\gamma}\right)^\lambda \left(\frac{\beta}{\gamma}\right)^{c-\lambda} \qquad (\gamma = \alpha + \beta)$$

zu, mit welcher zu erwarten war, daß aus der Hilfsurne λ-mal weiß und $c - \lambda$-mal schwarz gezogen werde.

Hiernach ist bei diesem Stande des Wissens

$$P_c = \frac{\omega_c p_c}{\sum\limits_1^c \omega_\lambda p_\lambda} = \frac{\alpha^c c^s}{\binom{c}{1} \alpha \beta^{c-1} 1^s + \binom{c}{2} \alpha^2 \beta^{c-2} 2^s + \cdots + \binom{c}{c} \alpha^c c^s},$$

und

$$P = \frac{\sum\limits_{\varkappa}^{c} \omega_\lambda p_\lambda}{\sum\limits_1^c \omega_\lambda p_\lambda},$$

wo $\varkappa = \frac{c}{2} + 1$ oder $\frac{c+1}{2}$, je nachdem c gerad oder ungerad ist.

Der Erkenntniswert dieser Wahrscheinlichkeiten ist größer als im vorigen Falle, weil sie sich auf ein umfangreicheres Wissen gründen.

Mit den früheren Werten $c = 5$ und $s = 4$ und mit $\alpha = 2$, $\beta = 1$ findet man:

$$P_5 = \frac{2^5 \cdot 5^4}{5 \cdot 2 \cdot 1^4 + 10 \cdot 2^2 \cdot 2^4 + 10 \cdot 2^3 \cdot 3^4 + 5 \cdot 2^4 \cdot 4^4 + 2^5 \cdot 5^4} = \frac{2000}{4761} = 0{,}4201$$

$$P = \frac{4696}{4761} = 0{,}9863.$$

103. Theorem von Bayes bei einer unbegrenzten Menge möglicher Ursachen.

Diejenige Form, in welcher die Frage nach der Wahrscheinlichkeit einer speziellen Ursache oder eines Ursachenkomplexes für ein beobachtetes Ereignis in den Anwendungen am häufigsten sich einstellt, läßt sich aus dem folgenden Schema entnehmen.

„Der beobachtete Erfolg E bestehe in dem m-maligen Eintreffen und dem n-maligen Ausbleiben eines Ereignisses $\mathfrak{E}$ in $s = m + n$ Versuchen, dem von vornherein jede Wahrscheinlichkeit zwischen 0 und 1 zugeschrieben werden kann. Über die Umstände, welche auf diese Wahrscheinlichkeit Einfluß haben, sei entweder gar nichts bekannt, so daß man genötigt ist, allen Werten aus dem bezeichneten Intervall gleichen Möglichkeitsgrad beizulegen, oder aber es bestehe nach dieser Richtung ein solches Wissen, daß man den Wahrscheinlichkeitsgrad jedes Wertes a priori anzugeben imstande ist."

Die Annahme eines Wertes x für die Wahrscheinlichkeit von $\mathfrak{E}$ ist hier als eine Ursache des beobachteten Ereignisses aufzufassen; es gibt also der Ursachen eine unendliche, nicht abzählbare Menge, weil die Menge der reellen Werte in einem Intervall von der Mächtigkeit eines Kontinuums ist. Mit der Supponierung eines Wertes x ist aber eine Hypothese über die Natur der zugrunde liegenden Umstände in der Regel nicht gemacht; nur ein allgemeines Bild ist hierfür ge-

schaffen, indem man sagen kann, es verhalte sich mit dem Ereignis in bezug auf seine Verwirklichungsmöglichkeit so, wie mit dem Ziehen einer weißen Kugel aus einer Urne, die mit einer unendlichen Menge weißer und schwarzer Kugeln in dem durch die Zahl x gekennzeichneten Mischungsverhältnis gefüllt ist. Die Menge der Kugeln muß als unendlich vorausgesetzt werden, damit das Mischungsverhältnis jedes Wertes fähig sei.

Für das beobachtete Ereignis entspringt aus x die Wahrscheinlichkeit

$$\mathfrak{W}_x(E) = \mathfrak{W}_x(\mathfrak{E}^m \bar{\mathfrak{E}}^n) = y = x^m (1 - x)^n;$$

mithin wäre bei gleicher apriorischer Wahrscheinlichkeit aller Werte von x die aposteriorische Wahrscheinlichkeit des besonderen Wertes x gleich

$$\frac{y}{\sum\limits_0^1 y}.$$

Aber ein solcher Ausdruck ist unausführbar, da die Summe im Nenner sich über *alle* Werte des Intervalles (0, 1) zu erstrecken hätte; er zeigt nur die von vornherein erkennbare Tatsache an, daß die Wahrscheinlichkeit eines individuellen Wertes x von Null nicht zu unterscheiden ist.

Hier ist also eine Modifikation der Fragestellung notwendig, und diese soll darin bestehen, daß man nicht um die Wahrscheinlichkeit eines einzelnen Wertes, sondern um die Wahrscheinlichkeit p_x eines Wertes aus dem Intervall $(x, x + dx)$ fragt; diese wird der Größe dx des Intervalls und dem zugehörigen Werte von y, der in dem Intervall als konstant erachtet werden kann, proportional, also durch

$$\varkappa y \, dx$$

darstellbar sein; bemerkt man, daß die auf alle Intervalle bezogene Summe dieser Ausdrücke notwendig den Wert 1 hat, weil ein Wert existent sein muß, so folgt aus der Gleichung

$$\varkappa \int_0^1 y \, dx = 1$$

für $\varkappa$ die Bestimmung:

$$\varkappa = \frac{1}{\int\limits_0^1 y \, dx},$$

und hiermit ergibt sich für die beschriebene Wahrscheinlichkeit der Ausdruck:

$$p_x = \frac{y\,dx}{\int\limits_0^1 y\,dx}. \tag{4}$$

Die Wahrscheinlichkeit, daß x in ein endliches Intervall (θ, θ') falle, ist daraus

$$P = \frac{\int\limits_\theta^{\theta'} y\,dx}{\int\limits_0^1 y\,dx}. \tag{5}$$

In dem Falle, daß man die Wahrscheinlichkeit a priori eines Wertes x anzugeben imstande ist — sie heiße w und wird eine Funktion von x sein — erfahren die Formeln (4), (5) eine ähnliche Änderung, wie sie sich bei dem Übergange von (1) zu (2) ergeben hat, und gehen über in:

$$p_x = \frac{w\,y\,dx}{\int\limits_0^1 w\,y\,dx}, \tag{6}$$

beziehungsweise

$$P = \frac{\int\limits_\theta^{\theta'} w\,y\,dx}{\int\limits_0^1 w\,y\,dx}. \tag{7}$$

Will man, um die Formeln (4) bis (7) unmittelbar verständlich zu machen, die Bedeutung der darin auftretenden Buchstaben symbolisch zum Ausdruck bringen, so bedarf es nur der Vereinbarung, daß unter $\mathfrak{W}(a, b)$ die Wahrscheinlichkeit gemeint sein soll, es falle x in das Intervall (a, b). Alsdann läßt sich das Ergebnis dieser Nummer in die Formeln zusammenfassen:

$$\mathfrak{W}_E(x, x + dx) = \frac{\mathfrak{W}_x(E)\,dx}{\int\limits_0^1 \mathfrak{W}_x(E)\,dx}, \tag{4*}$$

$$\mathfrak{W}(\theta, \theta') = \frac{\int\limits_\theta^{\theta'} \mathfrak{W}_x(E)\,dx}{\int\limits_0^1 \mathfrak{W}_x(E)\,dx}, \tag{5*}$$

$$\mathfrak{W}_E(x, x + dx) = \frac{\mathfrak{W}(x)\,\mathfrak{W}_x(E)\,dx}{\int\limits_0^1 \mathfrak{W}(x)\,\mathfrak{W}_x(E)\,dx}, \tag{6*}$$

$$\mathfrak{W}(\theta, \theta') = \frac{\int\limits_\theta^{\theta'} \mathfrak{W}(x)\,\mathfrak{W}_x(E)\,dx}{\int\limits_0^1 \mathfrak{W}(x)\,\mathfrak{W}_x(E)\,dx}. \tag{7*}$$

104. Die wahrscheinlichste Hypothese. Die wahrscheinlichste Hypothese über die bedingenden Umstände des beobachteten Ereignisses ist durch jenen Wert von x gekennzeichnet, welcher p_x, d. h. welcher y, beziehungsweise wy zum Maximum macht. Man nennt diesen Wert — er heiße a — die *Wahrscheinlichkeit von* $\mathfrak{E}$ *nach der wahrscheinlichsten Hypothese.*

Aus $y = x^m(1 - x)^n$ ergibt sich durch Nullsetzen des Differentialquotienten die Bestimmung

$$a = \frac{m}{m + n} = \frac{m}{s};$$

hiernach berechnet sich der maximale Wert von y:

$$M = \frac{m^m n^n}{s^s}.$$

Wir wollen nun das Verhalten von y in der Umgebung des Maximums unter der Voraussetzung prüfen, daß s eine große Zahl ist, und setzen zu diesem Ende

$$x = \frac{m}{s} + z,$$

wodurch

$$1 - x = \frac{n}{s} - z$$

und

$$y = \left(\frac{m}{s} + z\right)^m \left(\frac{n}{s} - z\right)^n = M\left(1 + \frac{s}{m} z\right)^m \left(1 - \frac{s}{n} z\right)^n = MZ$$

wird. Je kleiner z, um so eher darf die Entwicklung

$$\begin{aligned} l \cdot Z &= ml \cdot \left(1 + \frac{s}{m} z\right) + nl \cdot \left(1 - \frac{s}{n} z\right) \\ &= m\left(\frac{s}{m} z - \frac{s^2}{2m^2} z^2 + \cdots\right) + n\left(-\frac{s}{n} z - \frac{s^2}{2n^2} z^2 - \cdots\right) \end{aligned}$$

bei der zweiten Potenz dieses Arguments abgebrochen werden; es wird dann

$$l \cdot Z = -\frac{s^3}{2mn} z^2$$

und somit

$$Z = e^{-\frac{s^3}{2mn}z^2};$$

bei Berücksichtigung der Glieder mit z^3 träte im Exponenten das Glied $\frac{(n-m)s^4}{3m^2n^2}z^3$ hinzu, das im Vergleich zu dem beibehaltenen auf den Wert der Exponentialgröße nur unerheblichen Einfluß nimmt.

Der Ansatz

$$y = Me^{-\frac{s^3}{2mn}z^2} \tag{8}$$

zeigt nun, daß y in der Umgebung von a anfangs langsam, bald aber sehr rasch abnimmt und Werte erlangt, die im Vergleich zu M außerordentlich gering sind. Zur Erläuterung dieser Angaben möge folgendes dienen: Für $m = 300$, $n = 200$, $(s = 500)$ ist $a = \frac{3}{5}$, und es findet sich, daß

$$\text{für } z = \pm\frac{1}{100} \quad y = 0{,}90107\,M,$$

$$\text{„ „} = \pm\frac{5}{100} \quad \text{„} = 0{,}073965\,M,$$

$$\text{„ „} = \pm\frac{1}{10} \quad \text{„} = 0{,}000029929\,M,$$

$$\text{„ „} = \pm\frac{1}{5} \quad \text{„} = 0{,}0000000000000000080241\,M.$$

Wenn auch die Rechnung bei so großem z nicht mehr scharf genug ist, so zeigt sie doch, daß Annahmen über x, die nur einigermaßen von der wahrscheinlichsten Annahme a abweichen, im Vergleich zu dieser eine überaus geringe Wahrscheinlichkeit für sich haben.

Was den Maximalwert M von y selbst betrifft, so ist er außerordentlich klein und im vorliegenden Falle erst an der 147. Dezimalstelle mit einer bedeutsamen Ziffer besetzt.

Mit diesem Verhalten steht die Funktion y in einem gewissen Gegensatze zu der Funktion $x^n e^{-x}$, welche wir in Nr. 12 unter der Voraussetzung eines großen n betrachtet haben.

105. Beispiel XLIX. *In eine Urne wurden c Kugeln, weiße und schwarze, durch Auslosung mit einer Münze eingebracht: so oft Wappen fiel, wurde eine weiße, so oft Schrift fiel, eine schwarze Kugel eingelegt. Eine darauffolgende Ausführung von s Ziehungen, wobei die gezogene Kugel zurückgelegt worden, ergab m weiße und n schwarze Kugeln. Welches ist die wahrscheinlichste Hypothese über die Zusammensetzung der Urne?*

Vorausgesetzt wird, daß c und s große Zahlen bedeuten.

Unabhängig von jeder Rechnung können hier gewisse Erwägungen angestellt werden. Auf Grund des Entstehungsmodus der Urne, welcher

das vorgängige Wissen darstellt, wäre $\frac{1}{2}$ die wahrscheinlichste Annahme über die Wahrscheinlichkeit für das Erscheinen einer weißen Kugel. Aus der gemachten Erfahrung allein ergäbe sich hierfür $\frac{m}{s}$. Sollen beide Momente bei der Feststellung der wahrscheinlichsten Hypothese zusammenwirken, wie das die Natur der Sache fordert, so wird ein Wert zwischen $\frac{1}{2}$ und $\frac{m}{s}$ zu nehmen sein, falls nicht vielleicht $\frac{m}{s} = \frac{1}{2}$ ist. Bei der Entscheidung der Frage, welchem von den beiden Werten man näherrücken soll, wäre auf die Umfänge der Versuchsreihen Rücksicht zu nehmen, durch welche einerseits die Füllung der Urne, andrerseits die Beobachtungsreihe zustande kam. Niemand würde Anstand nehmen, wenn beide Reihen gleich umfangreich waren ($c = s$), den wahrscheinlichsten Wert in die Mitte zwischen $\frac{1}{2}$ und $\frac{m}{s}$ zu verlegen; ein Überwiegen des c würde das Vertrauen zu $\frac{1}{2}$, ein Überwiegen des s das Vertrauen zu $\frac{m}{s}$ erhöhen und ein Näherrücken an diesen Wert rechtfertigen.

Sehen wir nun zu, welche Antwort die Rechnung gibt.

Die Wahrscheinlichkeit, daß in die Urne statt $\frac{c}{2}$ weiße Kugeln deren $\frac{c}{2} - l$ kommen, ist (s. Nr. 75) proportional

$$e^{-\frac{2l^2}{c}};$$

ist dem so, dann ist $\frac{1}{2} - \frac{l}{c}$ die Wahrscheinlichkeit für das Ziehen einer weißen, $\frac{1}{2} + \frac{l}{c}$ die Wahrscheinlichkeit für das Ziehen einer schwarzen Kugel und

$$\left(\frac{1}{2} - \frac{l}{c}\right)^m \left(\frac{1}{2} + \frac{l}{c}\right)^n$$

die Wahrscheinlichkeit des beobachteten Ereignisses. Setzt man $\frac{l}{c} = z$, so entspricht

$$\frac{1}{2} - z$$

der wahrscheinlichsten Hypothese, wenn für z derjenige Wert gesetzt wird, der

$$e^{-2cz^2}\left(\frac{1}{2} - z\right)^m \cdot \left(\frac{1}{2} + z\right)^n$$

zu einem Maximum macht. Die Bedingung dafür ist:

$$-4cz - \frac{m}{\frac{1}{2} - z} + \frac{n}{\frac{1}{2} + z} = 0;$$

sie kann, da z nur als ein kleiner echter Bruch zu vermuten ist, angenähert durch die folgende ersetzt werden:

$$-4cz - 2m(1 + 2z) + 2n(1 - 2z) = 0;$$

daraus berechnet sich

$$z = \frac{n - m}{2(c + s)}$$

und

$$\frac{1}{2} - z = \frac{c + 2m}{2(c + s)}.$$

Dies ist aber das arithmetische Mittel aus $\frac{1}{2}$ und $\frac{m}{s}$, wenn man diesen Zahlen die Gewichte c und s beilegt; denn

$$\frac{c \cdot \frac{1}{2} + s \cdot \frac{m}{s}}{c + s} = \frac{c + 2m}{2(c + s)}.$$

Die Rechnung bestätigt also die allgemeinen Erwägungen vollinhaltlich.

106. Umkehrung des Bernoullischen Theorems. Während das Bernoullische Theorem die Erwartungsbildung für eine große Anzahl vorzunehmender Beobachtungen oder Versuche regelt, wenn die Wahrscheinlichkeiten für die dabei in Betracht kommenden, einander ausschließenden Ereignisse E, $\bar{E}$ bekannt sind und durch die Dauer der Versuche unverändert bleiben; so gestattet umgekehrt das Theorem von Bayes unter gewissen Voraussetzungen, aus dem Ergebnis einer ausgedehnten Beobachtungsreihe einen Schluß zu ziehen auf die unbekannten Wahrscheinlichkeiten der beteiligten Ereignisse; selbstverständlich kann dieser Schluß nur in einem Wahrscheinlichkeitsurteil bestehen. Vorausgesetzt wird dabei Unveränderlichkeit der bedingenden Umstände während der Beobachtungen und gleiche apriorische Wahrscheinlichkeit der möglichen Hypothesen.

Der beobachtete Erfolg bestehe in dem m-maligen Eintreffen eines Ereignisses E und in dem n-maligen Eintreffen des entgegengesetzten Ereignisses $\bar{E}$ in $m + n = s$ Versuchen; die unbekannten, zur Einheit sich ergänzenden Wahrscheinlichkeiten der Ereignisse seien p, q.

Die Wahrscheinlichkeit, daß p zwischen die Grenzen $\frac{m}{s} - \theta$ und $\frac{m}{s} + \theta$ falle, von seinem wahrscheinlichsten Werte $\frac{m}{s}$ also höchstens

um θ nach der einen oder andern Seite abweiche, ist nach Formel (5), Nr. 103, gleich

$$P = \frac{\int\limits_{\frac{m}{s}-\theta}^{\frac{m}{s}+\theta} y\,dx}{\int\limits_0^1 y\,dx},$$

darin $y = x^m(1-x)^n$ gesetzt.

Für das Integral im Nenner ergibt sich durch Anwendung partieller Integration leicht:

$$\int\limits_0^1 x^m(1-x)^n dx = \frac{m!\,n!}{(m+n+1)!}. \tag{9}$$

In dem Zählerintegral setze man

$$x = \frac{m}{s} + z;$$

es geht dadurch über in

$$\int\limits_{-\theta}^{\theta} \left(\frac{m}{s}+z\right)^m \left(\frac{n}{s}-z\right)^n dz,$$

wofür unter der Voraussetzung, daß die Grenzen des Integrals genügend eng seien, nach den Entwicklungen der Nr. 104 mit großer Annäherung

$$\frac{m^m n^n}{s^s} \int\limits_{-\theta}^{\theta} e^{-\frac{s^3 z^2}{2mn}} dz$$

geschrieben werden kann.

Mit dieser Umgestaltung und Approximation wird

$$P = \frac{(s+1)!}{m!\,n!} \frac{m^m n^n}{s^s} \int\limits_{-\theta}^{\theta} e^{-\frac{s^3 z^2}{2mn}} dz.$$

Wendet man auf die Fakultäten die Stirlingsche Formel an, so ergibt sich:

$$\frac{(s+1)!}{m!\,n!} = (s+1) \frac{s^s e^{-s}\sqrt{2\pi s}}{m^m n^n e^{-s} 2\pi\sqrt{mn}} = \frac{(s+1)s^s\sqrt{s}}{m^m n^n \sqrt{2\pi mn}},$$

wofür, je größer s, mit um so größerer Annäherung

$$\frac{s^s}{m^m n^n} \sqrt{\frac{s^3}{2\pi m n}}$$

substituiert werden kann.

Dann aber lautet P wie folgt:

$$P = \sqrt{\frac{s^3}{2\pi m n}} \int_{-\theta}^{\theta} e^{-\frac{s^3 z^2}{2mn}} dz;$$

führt man hierin die neue Variable t ein,

$$\sqrt{\frac{s^3}{2mn}}\, z = t$$

setzend, und bezeichnet $\theta \sqrt{\frac{s^3}{2mn}}$ mit γ, so gelangt man zu dem Satze:

„Es ist mit der Wahrscheinlichkeit

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-t^2} dt$$

zu erwarten, daß die unbekannte Wahrscheinlichkeit p des Ereignisses E zwischen den Grenzen

$$\frac{m}{s} - \gamma \sqrt{\frac{2mn}{s^3}} \quad \text{und} \quad \frac{m}{s} + \gamma \sqrt{\frac{2mn}{s^3}} \tag{11}$$

liege."

Da das Intervall der Grenzen, $2\gamma \sqrt{\frac{2mn}{s^3}}$, von der Ordnung $\frac{1}{\sqrt{s}}$ ist, weil m, n im allgemeinen von der Ordnung der Zahl s sind, so konvergiert es, wie groß auch γ, d. h. wie nahe auch P der Einheit angenommen wird, mit wachsendem s gegen Null. Man kann also die Zahl s der Beobachtungen so groß wählen, daß mit einer der Einheit beliebig nahen Wahrscheinlichkeit erwartet werden darf, die unbekannte Wahrscheinlichkeit p[1]) liege innerhalb beliebig eng festgesetzter Grenzen. In diesem Ausspruche liegt die *Umkehrung des Gesetzes der großen Zahlen.*

Sowie unter der Voraussetzung, daß die Wahrscheinlichkeit p von E *bekannt* sei, von wahrscheinlichen Grenzen der relativen Häufigkeit $\frac{m}{s}$ von E in s Versuchen und von einer Präzision der Versuchs-

1) Gleichgültig, ob es eine elementare oder eine Durchschnittswahrscheinlichkeit im eigentlichen Sinne ist. (Vgl. Nr. 97.)

reihe gesprochen wurde, so kann auch hier von wahrscheinlichen Grenzen der *unbekannten* Wahrscheinlichkeit p von E und von der Präzision ihrer Bestimmung durch die Versuchsreihe die Rede sein. Es sind nämlich im Sinne der Ausführungen von Nr. 79

$$\frac{m}{s} - 0{,}476936\sqrt{\frac{2mn}{s^3}} \quad \text{und} \quad \frac{m}{s} + 0{,}476936\sqrt{\frac{2mn}{s^3}}$$

die wahrscheinlichen Grenzen von p, d. h. die Grenzen, innerhalb deren sein Wert mit der Wahrscheinlichkeit $P = \frac{1}{2}$ zu erwarten ist, und nach Nr. 85 ist

$$h = \sqrt{\frac{s^3}{2mn}}$$

die Präzision der durch die Erfahrung gewonnenen Bestimmung $\frac{m}{s}$ für p.

Zur Ausführung der hier auftretenden Rechnungen bedient man sich der Tafel I.

Es handle sich beispielsweise um die folgende Frage: Aus einer Urne, welche 1 000 000 Kugeln enthält, weiße und schwarze in unbekanntem Mischungsverhältnis, sind 800 Kugeln gezogen worden[1]); davon waren 320 weiß und 480 schwarz; wie groß ist die Wahrscheinlichkeit, daß das Verhältnis der Anzahl der weißen Kugeln zu ihrer Gesamtzahl, dessen wahrscheinlichster Wert auf Grund der Beobachtung $\frac{2}{5}$ ist, zwischen die Grenzen $\frac{15}{40}$ und $\frac{17}{40}$ falle?

Wiewohl hier x, die hypothetische Wahrscheinlichkeit für das Ziehen einer weißen Kugel, nicht *aller* Werte zwischen 0 und 1 fähig ist, so sind die möglichen Werte doch so dicht, daß man bei der Fiktion, x sei eine stetige Variable, verbleiben darf.

Aus dem Ansatz

$$\gamma\sqrt{\frac{2 \cdot 320 \cdot 480}{800^3}} = \frac{1}{40}$$

berechnet sich $\gamma = 1{,}0205$; somit ist

$$P = \Phi(1{,}0205) = 0{,}85104.$$

Mit dieser Wahrscheinlichkeit ist anzunehmen, daß die Anzahl der weißen Kugeln in der Urne nicht unter 375 000 und nicht über 425 000 betrage.

107. Allgemeine Bemerkungen über die Wahrscheinlichkeiten von Ursachen. Man kann die Bedeutung einer a priori be-

---

1) Für das numerische Endresultat wird es hier gleichgültig sein, ob man sich die Kugeln zurückgelegt denkt oder nicht.

stimmten Wahrscheinlichkeit mittels des Gesetzes der großen Zahlen illustrieren, indem man sagt, in einer großen Anzahl von Versuchen werde sich das Ereignis höchstwahrscheinlich nahezu in der durch seine Wahrscheinlichkeit ausgedrückten relativen Häufigkeit zutragen. Diese Deutung haben einige Philosophen[1]) geradezu als Grundlage für die Definition der Wahrscheinlichkeit genommen, mitunter allerdings von einer prinzipiell mißverständlichen Auffassung des Bernoullischen Theorems ausgehend. Zu einer gemeinverständlichen Erklärung einer numerischen Wahrscheinlichkeit ist diese Deutung am besten geeignet.

Es entsteht nun die Frage: Lassen die Wahrscheinlichkeiten der Ursachen eine solche Deutung immer zu, nicht etwa bloß formell, sondern in berechtigter Weise?

Wir wollen diese Frage an Beispielen erörtern.

Aus einer Urne, von der man bloß weiß, daß sie vier Kugeln, die schwarz oder weiß sind, enthält, sind vier Ziehungen (mit jedesmaligem Zurücklegen der Kugel) gemacht worden und ergaben dreimal weiß und einmal schwarz. Nach der Bayesschen Regel ist die Wahrscheinlichkeit, daß die Urne zwei weiße und zwei schwarze Kugeln enthalte,

$$\frac{\left(\frac{1}{2}\right)^3 \frac{1}{2}}{\left(\frac{3}{4}\right)^3 \frac{1}{4} + \left(\frac{1}{2}\right)^3 \frac{1}{2} + \left(\frac{1}{4}\right)^3 \frac{3}{4}} = \frac{8}{23}.$$

Läßt dieses Resultat in *berechtigter* Weise eine Deutung nach dem Gesetz der großen Zahlen zu? Darf man sagen: Wenn man sehr viele Urnen nimmt, deren jede vier Kugeln enthält, die nur weiß oder schwarz sein können; wenn man aus jeder vier Ziehungen in der beschriebenen Weise macht und dabei jene Urnen ausscheidet, welche dreimal weiß und einmal schwarz ergaben, so werden nahe $\frac{8}{23}$ von diesen Urnen mit zwei weißen und zwei schwarzen Kugeln gefüllt sein?

Für diesen Schluß besteht kein innerer Grund, da man nicht weiß, wie die Urnen entstanden, d. h. wie sie gefüllt worden sind. Er hätte Berechtigung nur dann, wenn man wüßte, daß bei der Füllung alle möglichen Kombinationen, nämlich drei weiß, eine schwarz; zwei weiß, zwei schwarz; eine weiß, drei schwarz, gleichmäßig berücksichtigt worden sind. Ist dieses Wissen nicht vorhanden, dann kann der Schluß irreführen.

Wir ändern nun dasselbe Beispiel dahin ab, daß wir als feststehend voraussetzen, die Urne sei durch Auslosung mit einer Münze entstanden derart, daß für Wappen weiß und für Schrift schwarz ein-

1) J. St. Mill, A System of Logic (1. Aufl. 1843); J. Venn, Logic of chance (1866).

gelegt wurde. Jetzt gibt die Bayessche Regel für dieselbe Füllung bei demselben beobachteten Ereignis die Wahrscheinlichkeit

$$\frac{6 \cdot \left(\frac{1}{2}\right)^3 \frac{1}{2}}{4 \cdot \left(\frac{3}{4}\right)^3 \frac{1}{4} + 6 \cdot \left(\frac{1}{2}\right)^3 \frac{1}{2} + 4 \cdot \left(\frac{1}{4}\right)^3 \cdot \frac{3}{4}} = \frac{4}{9}.$$

Dieses Resultat läßt eine objektiv begründete Deutung in dem obigen Sinne zu wie folgt: Wenn man eine sehr große Anzahl von Urnen in der beschriebenen Weise füllt; wenn man aus jeder vier Ziehungen macht und diejenigen Urnen beiseite stellt, aus welchen dreimal weiß und einmal schwarz zum Vorschein kam; so ist es sehr wahrscheinlich, daß nahe $\frac{4}{9}$ dieser Urnen mit zwei weißen und zwei schwarzen Kugeln gefüllt sein werden.

In den Fällen, wo das apriorische Wahrscheinlichkeitsgesetz der Ursachen bekannt ist, haben die nach der Bayesschen Regel ermittelten Wahrscheinlichkeiten dieselbe Berechtigung und Bedeutung, wie a priori bestimmte Wahrscheinlichkeiten.

Wo jedoch die Kenntnis dieses Gesetzes mangelt und man genötigt ist, den Hypothesen gleiche apriorische Wahrscheinlichkeit zuzuschreiben, sind die Resultate nicht kritiklos hinzunehmen. Die schablonenhafte Anwendung der Bayesschen Regel kann hier zu Ergebnissen führen, die keinen inneren Wert besitzen und gegen deren Anerkennung sich der gemeine Verstand sträubt.

Indessen kann die mehrerwähnte Regel auch dann zu brauchbaren Resultaten führen, wenn die Annahme gleicher apriorischer Möglichkeit der Hypothesen zweifellos unbegründet ist; dies trifft zu, wenn der Rechnung sehr umfangreiche Erfahrungsdaten zugrunde gelegt werden können, und erklärt sich aus dem Verhalten der Wahrscheinlichkeit eines aus vielen Ereignissen zusammengesetzten Erfolges, in ihrer außerordentlich raschen Abnahme, sobald man sich von ihrem Maximum nach der einen oder andern Seite entfernt, aber auch durch den Umstand, daß in den meisten Fällen innerhalb nicht zu weiter Grenzen eine Abstufung des Wahrscheinlichkeitsgrades der Ursachen a priori wirklich untunlich ist. Hier bildet die ausgebreitete Erfahrung trotz des Mangels vorgängiger Kenntnisse eine zureichend feste Grundlage für die Beurteilung des Wahrscheinlichkeitsgrades.

Als Beispiel einer mißverständlichen Anwendung der Bayesschen Regel kann die folgende Frage gelten: Mit einer Münze ist einmal geworfen worden, und es erschien Wappen; wie groß ist die Wahrscheinlichkeit, daß die Wappenseite gegenüber der Schriftseite begünstigt sei?

Hat man nämlich keinen Grund, aus der Beschaffenheit der Münze

eine Ungleichheit der beiden Münzseiten betreffs ihrer Realisierungsmöglichkeit zu vermuten, dann ist der Wahrscheinlichkeitsansatz $\frac{1}{2}$ für beide innerlich, objektiv, so stark motiviert, daß die gemachte „Erfahrung“ gar keinen Einfluß üben kann. Ist aber Grund zu einer solchen Vermutung vorhanden, so wird die eine Erfahrung gewiß nicht für zureichend befunden werden, das Maß der Vermutung zu bestimmen.

Stellte man sich auf den Standpunkt, daß mangels einer genaueren Einsicht alle Werte von 0 bis 1 für die Wahrscheinlichkeit x der Wappenseite als gleichmöglich anzusehen seien, so ergäbe die Bayessche Regel als Wahrscheinlichkeit für die Begünstigung dieser Münzseite

$$\frac{\int_{\frac{1}{2}}^{1} x\,dx}{\int_{0}^{1} x\,dx} = \frac{3}{4},$$

ein Resultat, das durch sich selbst die Unzulässigkeit einer solchen Rechnungsweise verrät. Sollte die eine Beobachtung eine so große Wahrscheinlichkeit für die erwähnte Hypothese begründen?

In der Tat entspricht die Rechnung keineswegs dem Stande des vorhandenen Wissens. Wenn auch die Gestaltung der Münzseiten (das ungleiche Relief) die Vermutung nahelegt, daß ihnen verschiedene Wahrscheinlichkeitsgrade zukommen dürften, so wird man doch keineswegs zugeben, es könnte die Wahrscheinlichkeit einer Seite ebenso leicht $\frac{1}{8}$ wie $\frac{3}{8}$ oder $\frac{5}{8}$ oder $\frac{15}{16}$ usw. betragen, wie das die Rechnung annimmt. Vielmehr wird der Sachverhalt der sein, daß man eine Abweichung innerhalb gewisser enger Grenzen für möglich, außerhalb derselben aber für ausgeschlossen hält; daß man ferner innerhalb dieser Grenzen eine Abstufung des Möglichkeitsgrades vorzunehmen sich außerstande fühlt. Zu einer begründeten Feststellung des Grenzintervalls fehlt es aber an jeglichem Behelf. Indessen zeigt die Probe mit einer annehmbaren Hypothese, wie erheblich sich das Resultat der Rechnung durch diese Auffassung ändert. Angenommen, man würde nach einer Besichtigung der Münze eine Abweichung bis zu $\frac{1}{16}$ für möglich halten, dann wäre das Intervall, innerhalb dessen x sich bewegen kann, durch $\frac{7}{16}$ und $\frac{9}{16}$ begrenzt, und die Wahrscheinlichkeit der Begünstigung der Wappenseite, erschlossen aus dem angestellten Versuch, betrüge

$$\frac{\int\limits_{\frac{8}{16}}^{\frac{9}{16}} x\,dx}{\int\limits_{\frac{7}{16}}^{\frac{9}{16}} x\,dx} = \frac{17}{32} = 0{,}531\cdots.$$

In diesem Resultate wird man nichts Befremdliches mehr erblicken können, wenn überhaupt eine so geringe Erfahrung als zureichende Grundlage für eine Rechnung erachtet wird.

Wären neun Versuche gemacht worden und hätten alle das Resultat „Wappen“ ergeben, so wäre dies schon ein erheblicher Grund, zu vermuten, die Wappenseite sei begünstigt; die erste Rechnungsweise gibt als Maß dieser Vermutung

$$\frac{\int\limits_{\frac{1}{2}}^{1} x^9\,dx}{\int\limits_{0}^{1} x^9\,dx} = \frac{1023}{1024} = 0{,}99903\cdots,$$

nach der zweiten ist es bloß

$$\frac{\int\limits_{\frac{8}{16}}^{\frac{9}{16}} x^9\,dx}{\int\limits_{\frac{7}{16}}^{\frac{9}{16}} x^9\,dx} = \frac{9^{10}-8^{10}}{9^{10}-7^{10}} = 0{,}753\cdots.$$

Nicht in den Ziffern, sondern in der mit dem gemeinen Verstande harmonierenden Tatsache liegt der erkenntnistheoretische Wert dieser Resultate, daß die erweiterte Erfahrung den Wahrscheinlichkeitsgrad der Vermutung erheblich vermehrt hat.

108. Beurteilung einer Abweichung eines beobachteten Erfolges vom erwartungsmäßigen. Der Erkenntniswert einer a priori angesetzten numerischen Wahrscheinlichkeit hängt von dem Grade der Sicherheit ab, mit welchem man von der Gleichmöglichkeit der Einzelfälle überzeugt ist. Der ideale Fall einer absoluten Sicherheit hierüber wird dort, wo es sich um eine physische Urteilsmaterie handelt, kaum jemals eintreten. Wäre er vorhanden, so bestünde auch kein Zweifel darüber, daß eine Abweichung, die ein

beobachteter Erfolg dem erwartungsmäßigen, d. h. dem wahrscheinlichsten gegenüber zeigt, und wäre sie noch so groß, dem Zufall, also den während der Beobachtung beständig variierenden Umständen zuzuschreiben ist.

Besteht jene absolute Sicherheit nicht, ist vielmehr für einen Zweifel an der Gleichmöglichkeit der Fälle Raum vorhanden, so überträgt sich dieser Zweifel auch auf den Ursprung einer wahrgenommenen Abweichung, und es tritt die Frage auf: Ist die Abweichung ein Werk des Zufalls oder ist sie in der Urteilsmaterie selbst begründet?

Es liegt in der Natur der Sache, daß auf eine solche Frage niemals eine dezidierte Antwort wird zu geben sein. Sie kann nur zu Vermutungen Anlaß geben, und auch diese werden mit Vorsicht zu fassen sein. Wie weit kann hier die Rechnung zu Hilfe genommen werden?

Vor allem ist zu bemerken, daß die beiden Sachverhalte, die in der gestellten Frage einander entgegengehalten sind, sich nicht gegenseitig ausschließen; man kann nicht sagen, *entweder* ist die beobachtete Abweichung ein Werk des Zufalls *oder* sie rührt von einem Verhalten der Materie her, das von dem vorausgesetzten abweichend ist; vielmehr wird sie fast ausnahmslos das Resultat des Zusammenwirkens beider Veranlassungen sein. Die Antwort wird daher auch nicht dahin lauten können, daß das eine von beiden eher anzunehmen sei als das andere.

Jede Abweichung, mag sie noch so groß und ihre Wahrscheinlichkeit im Vergleich zur Wahrscheinlichkeit der Abweichung Null noch so klein sein, liegt im Bereich der Möglichkeit und hat daher an sich nichts Befremdendes. Aus *einer* Abweichung allein Schlüsse zu ziehen, wenn nicht schon a priori zu Vermutungen eine Veranlassung vorlag, hat logisch keine Berechtigung. Für eine wohlbegründete Rechnung fehlt es an den erforderlichen Daten.

Wenn trotzdem Rechnungen in solchen Fragen ausgeführt worden sind, so konnte ihren Resultaten eine entscheidende Bedeutung nicht beigemessen werden: es konnte sich vielmehr dabei nur darum handeln, der schätzungsweisen Vermutung einigen Halt zu gewähren.

Zur Erläuterung nehmen wir die von Poisson[1]) untersuchten Münzversuche Buffons auf.

Mit einer Münze sind 4040 Würfe gemacht worden; 2048-mal — statt 2020-mal, wie es dem wahrscheinlichsten Falle entsprechen würde — ist Wappen erschienen. Was kann daraus in bezug auf die Beschaffenheit der Münze vermutet werden?

---

1) Recherches etc., deutsche Bearbeitung p. 198 ff. Vgl. auch J. Bertrand, Calc. d. prob., p. 157 sq.

Zunächst könnte man sich die Frage vorlegen, ob denn die beobachtete Abweichung 28 im Zusammenhalt mit der Versuchszahl exorbitant zu nennen sei; und um hierfür einen Anhalt zu haben, könnte man die Wahrscheinlichkeit rechnen, mit welcher eine Abweichung von 28 oder darüber zu erwarten ist.

Aus dem Ansatze

$$\gamma \sqrt{2 \cdot 4040 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 28$$

berechnet sich

$$\gamma = 0{,}623;$$

mithin ist

$$P = \Phi(0{,}623) = 0{,}6217$$

die Wahrscheinlichkeit einer Abweichung bis zum Betrage 28,

$$1 - P = 0{,}3783$$

die Wahrscheinlichkeit einer Abweichung von 28 aufwärts. Man kann also eine Abweichung von 28 und darüber nicht als etwas Außergewöhnliches, höchst Unwahrscheinliches bezeichnen. Unter 1000 Versuchsreihen von je 4040 Würfen, ausgeführt mit *exakten* Münzen, hätte man gegen 378-mal eine Abweichung zu gewärtigen, die über 28 hinausgeht.

Man könnte sich ferner auf den Standpunkt stellen, daß man über die „Münze“ gar nichts wisse, und nun auf Grund der gemachten Beobachtung allein die Wahrscheinlichkeit suchen, welche der Hypothese zukommt, die Wappenseite sei gegenüber der Schriftseite begünstigt. Dies führt zu folgender Rechnung.

Bezeichnet man die Wahrscheinlichkeit der Wappenseite mit $\frac{1}{2} + z$, die der Schriftseite demgemäß mit $\frac{1}{2} - z$, so handelt es sich um die Wahrscheinlichkeit, daß $1 > \frac{1}{2} + z > \frac{1}{2} - z$, daß also z zwischen den Grenzen 0 und $\frac{1}{2}$ enthalten sei; für diese ergibt sich auf Grund der vorliegenden Beobachtung der Ausdruck:

$$\frac{\int_0^{\frac{1}{2}} \left(\frac{1}{2} + z\right)^{2048} \left(\frac{1}{2} - z\right)^{1992} dz}{\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{1}{2} + z\right)^{2048} \left(\frac{1}{2} - z\right)^{1992} dz}.$$

Seine strenge Auswertung wäre wegen der großen Exponenten beschwerlich. Wir suchen daher einen Näherungswert für die Funktion

Z unter den beiden Integralzeichen, indem wir ihren Logarithmus wie folgt entwickeln:

$$l \cdot Z = l \cdot \frac{1}{2^{4040}} + 2048\, l \cdot (1 + 2z) + 1992\, l \cdot (1 - 2z)$$

$$= l \cdot \frac{1}{2^{4040}} + 112z - 8080z^2 - \cdots.$$

Die Natur von Z bringt es mit sich, daß es zu beiden Seiten seines Maximums, das sich für $z = 0{,}00693$ ergibt, sehr bald außerordentlich rasch abnimmt, so daß nur ein enges Intervall um diesen besonderen Wert von z auf den numerischen Betrag von P, soweit er von Interesse ist, Einfluß hat; dies ist auch der Grund, warum man die Entwicklung mit der zweiten Potenz abbrechen kann. Durch den Übergang zur Zahl selbst erhält man:

$$Z = \frac{1}{2^{4040}} e^{112z - 8080z^2}$$

und daraus:

$$P = \frac{\int\limits_0^{\frac{1}{2}} e^{112z - 8080z^2}\, dz}{\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} e^{112z - 8080z^2}\, dz}.$$

Während nun der Exponent an der Stelle $z = 0{,}00693$, wo die Funktion ihren größten Wert annimmt, $1{,}164 \cdots$ beträgt, sinkt er bis zur oberen Grenze auf den Betrag -1964 herab; infolgedessen hat es keinen Einfluß auf die maßgebenden Ziffern von P, wenn man das Integral bis ∞ erstreckt; ähnliches gilt bezüglich der unteren Grenze des Nennerintegrals, so daß man schreiben kann:

$$P = \frac{\int\limits_0^{\infty} e^{112z - 8080z^2}\, dz}{\int\limits_{-\infty}^{\infty} e^{112z - 8080z^2}\, dz}$$

$$= \frac{\int\limits_0^{\infty} e^{112z - 8080z^2}\, dz}{\int\limits_0^{\infty} e^{112z - 8080z^2}\, dz + \int\limits_0^{\infty} e^{-112z - 8080z^2}\, dz}.$$

Allgemein ist

$$\int_0^\infty e^{2bz-a^2z^2}dz = \int_0^\infty e^{\frac{b^2}{a^2}-\left(az-\frac{b}{a}\right)^2}dz = \frac{1}{a}e^{\frac{b^2}{a^2}}\int_{-\frac{b}{a}}^\infty e^{-t^2}dt$$

$$= \frac{1}{a}e^{\frac{b^2}{a^2}}\left\{\int_0^\infty e^{-t^2}dt + \int_0^{\frac{b}{a}} e^{-t^2}dt\right\}$$

$$= \frac{\sqrt{\pi}}{2a}e^{\frac{b^2}{a^2}}\left\{1 + \Phi\left(\frac{b}{a}\right)\right\}.$$

und analog ergibt sich

$$\int_0^\infty e^{-2bz-a^2z^2}dz = \frac{\sqrt{\pi}}{2a}e^{\frac{b^2}{a^2}}\left\{1 - \Phi\left(\frac{b}{a}\right)\right\}.$$

Durch Anwendung dieser Formeln auf den obigen Ausdruck erhält man mittels der Tafel I:

$$P = \frac{1 + \Phi(0{,}6229)}{2} = \frac{1{,}62162}{2} = 0{,}8108.$$

Diesem Resultate gegenüber hat der Einwand, daß es unter der Annahme entstanden ist, als seien alle Werte zwischen 0 und 1 für die Wahrscheinlichkeit einer Münzseite gleich zulässig, keine Berechtigung; in Wirklichkeit haben nämlich auf die abgeleiteten Ziffern nur die Werte eines engen Intervalls um $z = 0{,}00693$ Einfluß.

Wie kann das Resultat im Sinne des Gesetzes der großen Zahlen gedeutet werden?

Wenn man mit einer ungeheuren Zahl von Münzen verschiedensten Gepräges Versuche machte, mit jeder 4040; wenn man diejenigen ausschiede, bei welchen sich 2048-mal Wappen ergab, so wäre man berechtigt zu erwarten, daß auf je 1000 dieser Münzen nahe 810 solcher Exemplare kämen, bei denen die Wappenseite begünstigt ist; bei den übrigen 190 wäre jenes Resultat eingetroffen, obwohl sie exakt oder zugunsten der *Schriftseite* unregelmäßig sind.

Eine Verifikation dieses Schlusses aber ist, auch wenn es möglich wäre, die hinreichende Menge von Versuchsreihen anzustellen, untunlich; denn es gibt kein Mittel, die Wahrscheinlichkeit der beiden Seiten einer Münze durch ein physikalisches Verfahren zu bestimmen.

Man kann das Ergebnis der Rechnungen wohl nicht anders als dahin zusammenfassen, daß die beobachtete Abweichung auch bei einer exakten Münze nichts Ungewöhnliches bedeuten würde, daß sie aber auch der Vermutung einer Unregelmäßigkeit der Münze zugunsten der Wappenseite einen erheblichen Wahrscheinlichkeitsgrad verleiht.

§ 2. Wahrscheinlichkeit künftiger Ereignisse auf Grund von Beobachtungen.

109. Aposteriorische Wahrscheinlichkeit eines zu gewärtigenden Ereignisses. Eine aus Versuchen oder Beobachtungen für ein zufälliges Ereignis abgeleitete Wahrscheinlichkeit wird als *empirische* oder als *Erfahrungswahrscheinlichkeit* bezeichnet. Die Theorie bietet zwei Methoden ihrer Bestimmung dar.

Die eine Methode besteht darin, daß man dem Ereignis diejenige Wahrscheinlichkeit zuschreibt, welche sich aus der wahrscheinlichsten Hypothese über das beobachtete Ereignis dafür ergibt. Man nennt diese Wahrscheinlichkeit die *nach der wahrscheinlichsten Ursache* bestimmte. Ihre Berechnung erfordert die Anwendung der Bayesschen Regel.

Die andere Methode besteht darin, daß man alle Hypothesen, die mit dem beobachteten Ereignis vereinbar sind, bei der Bestimmung der Wahrscheinlichkeit des künftigen Ereignisses mitwirken läßt, jede entsprechend dem Grade ihrer eigenen Wahrscheinlichkeit. In diesem Falle spricht man von einer *Wahrscheinlichkeit a posteriori* im besonderen.

Vom theoretischen Standpunkte wäre die zweite Methode vorzuziehen, weil sie auf alle möglichen Ursachen Rücksicht nimmt, während die erste sich auf eine bestimmte unter ihnen allein stützt. Bei umfangreichen Erfahrungsreihen, wie solche in den praktischen Anwendungen zumeist vorliegen, liegen die Resultate beider Methoden so nahe beieinander, daß es wohl gleichgültig bleibt, welches von beiden man wählt. Zumeist wird hier unter der empirischen Wahrscheinlichkeit die nach der wahrscheinlichsten Hypothese gerechnete verstanden.

Das erste Verfahren bedarf keiner Erläuterung mehr.

Die Bestimmung der aposteriorischen Wahrscheinlichkeit besteht in einer Anwendung der Sätze von der zusammengesetzten und der vollständigen Wahrscheinlichkeit.

Sei E das beobachtete Ereignis, das sich aus den einfachen, einander entgegengesetzten Ereignissen $\mathfrak{E}$, $\bar{\mathfrak{E}}$ irgendwie zusammensetzt; dasselbe lasse die Ursachen $U_i (i = 1, 2, \cdots n)$ zu, deren apriorische Wahrscheinlichkeiten ω_i $(i = 1, 2, \cdots n)$ sein mögen; p_i sei die Wahrscheinlichkeit, mit welcher E aus der Ursache U_i zu erwarten ist, und P_i die aposteriorische Wahrscheinlichkeit dieser Ursache; endlich bezeichne $\mathfrak{p}_i$ die Wahrscheinlichkeit, welche die Ursache U_i dem künftigen (ebenfalls aus $\mathfrak{E}$ und $\bar{\mathfrak{E}}$ zusammengesetzten) Ereignisse F verleiht.

Dann ist

$$\Pi = P_1 \mathfrak{p}_1 + P_2 \mathfrak{p}_2 + \cdots + P_n \mathfrak{p}_n \tag{12}$$

die vollständige, aus der Erfahrung gefolgerte, also die Wahrschein-

lichkeit a posteriori von F; denn P_1 ist die aus der Erfahrung E abgeleitete Wahrscheinlichkeit, daß die Ursache U_1 vorlag, und wenn sie vorlag und wenn bei dem künftigen Ereignis die nämliche Ursache aktiv ist wie bei dem beobachteten, so ist F mit der Wahrscheinlichkeit $\mathfrak{p}_1$ zu erwarten; folglich ist $P_1 \mathfrak{p}_1$ die Wahrscheinlichkeit, daß F durch die Ursache U_1 zustande kommt usw.

Nach dem Theorem von Bayes (Nr. 101) ist

$$P_i = \frac{\omega_i p_i}{\sum_1^n \omega_i p_i},$$

folglich ist, durch die einfachsten Rechenelemente ausgedrückt:

$$\Pi = \frac{\sum_1^n \omega_i p_i \mathfrak{p}_i}{\sum_1^n \omega_i p_i}. \qquad (13)$$

Waren die Ursachen von vornherein gleichmöglich, also $\omega_1 = \omega_2 = \cdots \omega_n$, so vereinfacht sich die Formel und lautet:

$$\Pi = \frac{\sum_1^n p_i \mathfrak{p}_i}{\sum_1^n p_i}. \qquad (14)$$

Die ganze Darstellung gewinnt an Durchsichtigkeit, wenn man sich des Begriffs der relativen Wahrscheinlichkeit bedient. Man kann unmittelbar den Ansatz bilden:

$$\mathfrak{W}_E(F) = \mathfrak{W}_E(U_1 F) + \mathfrak{W}_E(U_2 F) + \cdots + \mathfrak{W}_E(U_n F),$$

der besagt, F könne mit einer der Ursachen zusammentreffen, immer unter der Voraussetzung, daß E eingetroffen sei. Nun ist allgemein

$$\mathfrak{W}_E(U_i F) = \mathfrak{W}_E(U_i)\mathfrak{W}_{U_i}(F)$$

und nach Formel (3*) in Nr. 101

$$\mathfrak{W}_E(U_i) = \frac{\mathfrak{W}(U_i)\mathfrak{W}_{U_i}(E)}{\sum_1^n {}^i\, \mathfrak{W}(U_i)\mathfrak{W}_{U_i}(E)},$$

daher endgültig

$$\mathfrak{W}_E(F) = \frac{\sum_1^n {}^i\, \mathfrak{W}(U_i)\mathfrak{W}_{U_i}(E)\mathfrak{W}_{U_i}(F)}{\sum_1^n {}^i\, \mathfrak{W}(U_i)\mathfrak{W}_{U_i}(E)}, \qquad (13^*)$$

übereinstimmend mit dem Inhalt der Formel (13).

Ist die Wahrscheinlichkeit x von $\mathfrak{E}$ aller Werte des stetigen Bereiches (0, 1) fähig, so stellen sich die früher mit ω, p, $\mathfrak{p}$ bezeichneten Wahrscheinlichkeiten insgesamt als Funktionen von x dar und sollen nunmehr mit w, y, z bezeichnet werden, so daß w die apriorische Wahrscheinlichkeit von x, y die auf x basierende Wahrscheinlichkeit von E und z die aus x gefolgerte Wahrscheinlichkeit von F bedeutet. Die aposteriorische Wahrscheinlichkeit von F stellt sich dann in der Gestalt

$$\Pi = \frac{\int_0^1 wyz\,dx}{\int_0^1 wy\,dx}, \tag{15}$$

und wenn w konstant ist, in der Form

$$\Pi = \frac{\int_0^1 yz\,dx}{\int_0^1 y\,dx} \tag{16}$$

dar.

Die Bemerkungen, die in Nr. 107 über die Wahrscheinlichkeit der Ursachen gemacht worden sind, übertragen sich notwendig auch auf Π.

Lediglich dazu, um den prinzipiellen Unterschied der beiden Bestimmungsweisen der empirischen Wahrscheinlichkeit hervorzuheben, diene das folgende einfache Beispiel (vgl. Nr. 107).

Aus einer Urne, die nur weiße und schwarze Kugeln, im ganzen vier, enthält, sind vier Ziehungen (mit Zurücklegung der Kugel) gemacht worden; dreimal erschien eine weiße, einmal eine schwarze Kugel. Welche empirische Wahrscheinlichkeit folgt aus diesen Tatsachen für das Ziehen einer weißen Kugel aus der Urne?

Dem beobachteten Ereignis E (3 weiß, 1 schwarz) können drei Ursachen zugrunde gelegt werden, nämlich:

U_1: 3 weiße, 1 schwarze Kugel,
U_2: 2 „ 2 „ Kugeln,
U_3: 1 „ 3 „ „ ;

die Wahrscheinlichkeiten von E nach diesen drei Ursachen sind:

$$p_1 = \left(\frac{3}{4}\right)^3 \frac{1}{4} = \frac{27}{256}, \quad p_2 = \left(\frac{1}{2}\right)^4 = \frac{16}{256}, \quad p_3 = \left(\frac{1}{4}\right)^3 \frac{3}{4} = \frac{3}{256};$$

die Wahrscheinlichkeiten der Ursachen selbst:

$$P_1 = \frac{27}{46}, \quad P_2 = \frac{16}{46}, \quad P_3 = \frac{3}{46}.$$

Das künftige Ereignis F ist das Ziehen einer weißen Kugel. Seine Wahrscheinlichkeit nach der wahrscheinlichsten Ursache (U_1) ist

$$\frac{3}{4} = 0{,}75,$$

seine Wahrscheinlichkeit a posteriori dagegen

$$\Pi = \frac{27}{46} \cdot \frac{3}{4} + \frac{16}{46} \cdot \frac{1}{2} + \frac{3}{46} \cdot \frac{1}{4} = \frac{29}{46} = 0{,}630 \cdots.$$

110. Hauptaufgabe über die aposteriorische Wahrscheinlichkeit. Das beobachtete Ereignis E bestehe in dem m-maligen Eintreffen und dem n-maligen Ausbleiben eines Ereignisses $\mathfrak{E}$ von unbekannter, a priori aller Werte gleichfähiger Wahrscheinlichkeit x. Das künftige Ereignis F bedeute das m'-malige Eintreffen und das n'-malige Ausbleiben von $\mathfrak{E}$ in $m' + n'$ weiteren Beobachtungen.

Die Bestimmung der aposteriorischen Wahrscheinlichkeit Π von F hat nach der Formel (16) zu erfolgen, und zwar ist darin:

$$y = x^m (1 - x)^n, \quad z = \binom{m' + n'}{m'} x^{m'} (1 - x)^{n'}$$

zu setzen; denn das beobachtete Ereignis zeigt eine bestimmte Reihenfolge des Eintreffens und Ausbleibens von $\mathfrak{E}$, während bezüglich des künftigen diese Reihenfolge unbestimmt und gleichgültig ist; aber auch wenn man sich auf den Standpunkt stellen wollte, daß die erste Reihenfolge unbekannt sei, würde der Binomialkoeffizient $\binom{m+n}{m}$ aus der Rechnung fallen. Hiernach hat man

$$\Pi = \binom{m' + n'}{m'} \frac{\int_0^1 x^{m+m'} (1 - x)^{n+n'} dx}{\int_0^1 x^m (1 - x)^n dx}. \tag{17}$$

Werden die Integrale nach der Formel (9), Nr. 106, ausgeführt, so ergibt sich für Π der nur aus Fakultäten zusammengesetzte Ausdruck:

$$\Pi = \frac{(m' + n')!\,(m + m')!\,(n + n')!\,(m + n + 1)!}{m'!\,n'!\,(m + m' + n + n' + 1)!\,m!\,n!}, \tag{17*}$$

der bei großen Werten von m, n, m', n' mittels der Stirlingschen Formel auszuwerten wäre.

Durch die Annahme $m' = 1$, $n' = 0$ geht Π in die *aposteriorische Wahrscheinlichkeit des Ereignisses* $\mathfrak{E}$ selbst über; diese beträgt also:

$$\Pi = \frac{\int_0^1 x^{m+1} (1 - x)^n dx}{\int_0^1 x^m (1 - x)^n dx} = \frac{m + 1}{m + n + 2}. \tag{18}$$

Sie unterscheidet sich um so weniger von der Wahrscheinlichkeit nach der wahrscheinlichsten Hypothese, deren Wert $\frac{m}{m+n}$ ist, je umfangreicher die Erfahrung E ist.

Unter der Voraussetzung, daß $m+n$ eine sehr große Zahl ist, hat das Resultat (18) auch dann eine wohlbegründete Bedeutung, wenn die Voraussetzung der gleichen apriorischen Wahrscheinlichkeit aller Werte von x nicht zutrifft; wegen der außerordentlich raschen Abnahme der Funktionen unter den Integralzeichen, sobald man sich von den einander naheliegenden Stellen $\frac{m+1}{m+n+1}$ und $\frac{m}{m+n}$ ihrer Maxima entfernt, wirken nämlich nur enge Intervalle um diese Werte auf den Betrag von Π ein, und innerhalb dieser Intervalle wird es zumeist untunlich sein, eine Abstufung im Möglichkeitsgrade der Werte von x vorzunehmen.

Setzt man in der Formel (17) $m = m'$, $n = n'$, $[m+n=s]$, so kommt man zu der Wahrscheinlichkeit eines künftigen Erfolgs, der mit dem beobachteten, was die Wiederholungszahlen des Eintreffens und Ausbleibens anlangt, vollständig übereinstimmt; ist s sehr groß, so ergibt sich für diese Wahrscheinlichkeit der Näherungswert

$$\Pi = \sqrt{\frac{s}{4\pi m n}}; \tag{19}$$

bei demselben s ist dieses Π um so größer, je mehr die Zahlen m, n voneinander verschieden sind; sein kleinster Wert ist $\frac{1}{\sqrt{\pi s}}$. Hätte man beispielsweise aus einer Urne unbekannten Inhalts in 10000 Ziehungen 2000-mal eine weiße und 8000-mal eine schwarze Kugel hervorgeholt, so wäre gemäß der Formel (19) mit der Wahrscheinlichkeit $\frac{1}{80\sqrt{\pi}} = 0{,}0075024$ zu erwarten, daß 10000 weitere Ziehungen dasselbe Resultat hervorbringen.

Zu der Formel (19) ist folgendes zu bemerken. Wären $\frac{m}{s}, \frac{n}{s}$ die *bekannten* Wahrscheinlichkeiten für das Eintreffen und Ausbleiben von $\mathfrak{E}$, so bestünde nach Nr. 74 die Wahrscheinlichkeit

$$\frac{1}{\sqrt{2\pi s \frac{m}{s} \cdot \frac{n}{s}}} = \sqrt{\frac{s}{2\pi m n}}$$

dafür, daß in $s = m+n$ Versuchen $\mathfrak{E}$ m-mal eintreffen und n-mal ausbleiben werde. Diese Wahrscheinlichkeit ist $\sqrt{2}$-mal größer als die in Formel (19) unter andern Voraussetzungen berechnete. Der Grund dieser Abweichung liegt darin, daß dort $\frac{m}{s}, \frac{n}{s}$ nicht die *sicheren*,

sondern nur die wahrscheinlichsten Werte der Wahrscheinlichkeiten für $\mathfrak{E}$ und $\bar{\mathfrak{E}}$ sind.

111. Beispiel I. *Eine Urne enthält 8 Kugeln, weiße und schwarze; 4 davon sind nach und nach herausgenommen worden, und es waren 3 weiß, 1 schwarz. Wie groß ist die Wahrscheinlichkeit, daß von 3 weiteren Kugeln, die man sukzessive zieht, 1 weiß und 2 schwarz sein werden?*

Fünf Hypothesen sind mit der bekannten Anzahl der Kugeln und dem beobachteten Ereignis vereinbar, nämlich

$U_1 \ldots 3$ weiße, 5 schwarze Kugeln
$U_2 \ldots 4$ „ 4 „ „
$U_3 \ldots 5$ „ 3 „ „
$U_4 \ldots 6$ „ 2 „ „
$U_5 \ldots 7$ „ 1 „ Kugel.

Die aus ihnen entspringenden Wahrscheinlichkeiten des *beobachteten* Erfolges sind:

$$p_1 = \frac{3\cdot 2\cdot 1\cdot 5}{8\cdot 7\cdot 6\cdot 5} = \frac{5}{280}, \quad p_2 = \frac{4\cdot 3\cdot 2\cdot 4}{8\cdot 7\cdot 6\cdot 5} = \frac{16}{280}, \quad p_3 = \frac{5\cdot 4\cdot 3\cdot 3}{8\cdot 7\cdot 6\cdot 5} = \frac{30}{280},$$

$$p_4 = \frac{6\cdot 5\cdot 4\cdot 2}{8\cdot 7\cdot 6\cdot 5} = \frac{40}{280}, \quad p_5 = \frac{7\cdot 6\cdot 5\cdot 1}{8\cdot 7\cdot 6\cdot 5} = \frac{35}{280};$$

jene des *zukünftigen* Ereignisses:

$$\mathfrak{p}_1 = 0, \quad \mathfrak{p}_2 = 3\cdot\frac{1\cdot 3\cdot 2}{4\cdot 3\cdot 2} = \frac{3}{4}, \quad \mathfrak{p}_3 = 3\cdot\frac{2\cdot 2\cdot 1}{4\cdot 3\cdot 2} = \frac{1}{2}, \quad \mathfrak{p}_4 = 0, \quad \mathfrak{p}_5 = 0.$$

Folglich reduziert sich die aposteriorische Wahrscheinlichkeit des künftigen Ereignisses auf

$$\Pi = P_2\mathfrak{p}_2 + P_3\mathfrak{p}_3;$$

wenn man in Ermangelung weiteren Wissens die Ursachen als a priori gleichwahrscheinlich ansieht, so ist

$$P_2 = \frac{16}{5+16+30+40+35} = \frac{8}{63}, \quad P_3 = \frac{15}{63},$$

demnach

$$\Pi = \frac{8}{14}.$$

Nun nehme man an, dieselben Ziehungen würden aus einer Urne mit einer unbegrenzten Menge weißer und schwarzer Kugeln unbekannten Mischungsverhältnisses gemacht worden sein, und es handelte sich um die aposteriorische Wahrscheinlichkeit des nämlichen künftigen Ereignisses. Die unbekannte Wahrscheinlichkeit x für das Ziehen einer weißen Kugel kann dann jeden Wert zwischen 0 und 1

besitzen und bleibt während der Ziehungen konstant. Man hat nun Π nach der Formel (17) zu rechnen und findet

$$\Pi = \binom{3}{1} \frac{\int\limits_0^1 x^4(1-x)^3\,dx}{\int\limits_0^1 x^3(1-x)\,dx} = 3 \cdot \frac{\frac{4!\ 3!}{8!}}{\frac{3!\ 1!}{5!}} = \frac{3}{14}.$$

Das Resultat ist dasselbe wie vorhin und diese Übereinstimmung gilt für beliebige Zahlen.[1])

112. Empirische Wahrscheinlichkeitsbestimmung. Neben die Ermittlung der Wahrscheinlichkeit auf Grund einer Analyse der Urteilsmaterie tritt als selbständige Methode die Wahrscheinlichkeitsbestimmung auf Grund von Erfahrungsdaten oder Versuchen über dieselbe.[2])

Was den *Erkenntniswert* der nach diesen beiden Methoden gebildeten Wahrscheinlichkeiten anlangt, so läßt sich eine vergleichende Aussage von allgemeiner Gültigkeit darüber nicht machen. Bei der ersten Methode hängt er von dem Grade der Sicherheit ab, mit welchem man über das relative Maß der Möglichkeit der Einzelfälle zu urteilen imstande ist, insbesondere von dem Grade der Berechtigung, mit der man die Einzelfälle als gleichmöglich annehmen darf. Bei der zweiten Methode ist für den Erkenntniswert der Umfang der Erfahrungen und die richtige Wertung der einzelnen maßgebend, die bei der Bildung des Wahrscheinlichkeitsbruches in Rechnung zu ziehen ist; auch hier wird man zumeist die einzelnen Erfahrungen als gleichwertig erachten müssen in Ermanglung eines verläßlichen Mittels zu ihrer Abwägung. Es kann hiernach ganz wohl geschehen, daß man einer auf breiter Erfahrungsgrundlage ermittelten Wahrscheinlichkeit den Vorzug geben wird, selbst dann, wenn eine apriorische Wahrscheinlichkeitsbestimmung formell durchführbar, aber materiell nicht genügend sicher ist. So würden, um ein Beispiel anzuführen, die aus einer umfangreichen Versuchsreihe mit einem Würfel ermittelten Wahrscheinlichkeiten der sechs Seitenflächen eine größere Berechtigung beanspruchen dürfen als die aus der bis zu einem gewissen Grade immer unsicheren Annahme der Gleichmöglichkeit gezogene Bewertung mit $\frac{1}{6}$.

Was die *Tragweite* der beiden Methoden betrifft, so ist die empirische Wahrscheinlichkeitsbestimmung der apriorischen weitaus überlegen; sobald man nämlich das Gebiet der Wirklichkeit betritt, hört die Durchführbarkeit des letzteren Verfahrens fast völlig auf. Fragen,

---

1) J. Todhunter, History of the mathematical theorie of probability (1865), p. 454 ff.

2) R. Lämmel, l. c., p. 36 ff., nennt diese Methode die „statistische“.

die die Naturwissenschaften und das soziale Leben nach dieser Richtung stellen, sind zumeist so komplizierter Art, daß an eine Wahrscheinlichkeitsbestimmung a priori gar nicht gedacht werden kann. Aber auch bei Problemen, deren Struktur eine Analyse zulassen würde, kann die weitere Verfolgung an den mathematischen Schwierigkeiten scheitern, die sich einer Bewertung der ins Verhältnis zu setzenden Mengen entgegenstellen.

Einige Beispiele mögen diese Aussagen ins volle Licht rücken. Man kann die Frage nach der Wahrscheinlichkeit einer männlichen Geburt stellen; es soll zunächst davon abgesehen werden, ob dabei von einer konstanten Wahrscheinlichkeit gesprochen werden kann in dem Sinne, daß auf Grund des Bernoullischen Theorems danach die absolute und relative Häufigkeit männlicher Geburten in einem bestimmten Gebiet und während eines bestimmten Zeitraumes beurteilt werden könnte. Auf den ersten Blick ist die Undurchführbarkeit einer apriorischen Lösung zu erkennen, weil über die das Geschlecht bestimmenden Umstände nicht einmal qualitativ, geschweige denn quantitativ etwas ausgesagt werden kann; es bleibt nur der Weg der empirischen Bestimmung offen. Ganz ebenso verhält es sich mit der Frage nach der Wahrscheinlichkeit, daß eine Person bestimmten Alters und gegebener Zugehörigkeit zu einem Personenkreise innerhalb eines bestimmten Altersintervalls sterbe, oder daß eine männliche Person unter den gleichen Bestimmungen heirate, eine an einer bestimmten Krankheit erkrankte Person genese usw.

Als Beispiel eines Falles, wo die mathematischen Schwierigkeiten das Hindernis bilden, sei die Frage nach der Wahrscheinlichkeit angeführt, daß zwei unabhängig voneinander gewählte natürliche Zahlen teilerfremd seien oder ein mit beliebigen Zahlen hingeschriebener gemeiner Bruch sich nicht abkürzen lasse.[1]) Wiewohl es sich hier um einen Sachverhalt handelt, der nicht vom Zufall abhängt, sondern den Gesetzen der Zahlen unterworfen ist, so kann doch mit voller Berechtigung von einer Wahrscheinlichkeit gesprochen werden. In der Form, in der die Frage gestellt ist, muß an die transfinite Menge der Zahlenpaare gedacht werden; geht man bis zur endlichen Zahl z, so hat man es mit z^2 Zahlenpaaren zu tun, und befinden sich darunter $\varphi(z)$ solche, die relativ prim sind, so wäre die Aufgabe nach dem Ansatz

$$\lim_{z=\infty} \frac{\varphi(z)}{z^2}$$

zu lösen, der aber erst dann ausführbar wäre, wenn man eine Darstellung von $\varphi(z)$ besäße. In Ermanglung einer solchen bleibt nichts

1) R. Lämmel, l. c., p. 39.

übrig, als zum empirischen Verfahren zu greifen. Zu dem Zwecke empfiehlt es sich, die Zahlenpaare durch die Punkte eines quadratischen Gitters darzustellen, die zu relativ primen Paaren gehörigen Punkte hervorzuheben und dann in dem Gebiet, auf das man sich beschränken will, zu zählen, Fig. 16; bei dieser Zählung wird man von dem Umstande Gebrauch machen, daß das ganze Schema bezüglich der durch die Punkte z/z verlaufenden Achse symmetrisch ist. Eine bis $z = 44$ geführte Untersuchung ergab folgende Resultate [1]):

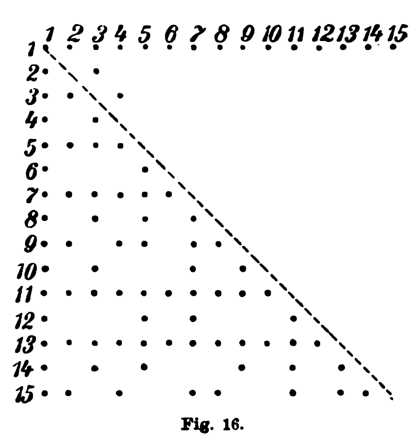


Fig. 16.

z	z^2	$\varphi(z)$	$\frac{(\varphi)z}{z^2}$
4	16	11	0,6625
8	64	43	0,67187
12	144	91	0,63194
16	256	159	0,62109
20	400	255	0,6375
24	576	359	0,62326
28	784	483	0,61607
32	1024	647	0,63183
36	1296	791	0,61034
40	1600	979	0,61187
44	1932	1207	0,62474

Wiewohl die Primzahlen immer wieder eine Störung des gleichmäßigen Verlaufs hervorrufen, ist auf Grund dieses Bruchstücks doch anzunehmen, daß der obige Grenzwert nahe an 0,6 liegen dürfte.

IV. Abschnitt. Bewertung von Vor- und Nachteilen, welche an zufällige Ereignisse geknüpft sind.

§ 1. Die mathematische Erwartung.

113. Definition der mathematischen Erwartung. Eine Person sei an einer Reihe zufälliger, einander ausschließender Ereignisse $F, F', \ldots F^{(n)}$ in der Weise interessiert, daß das Eintreffen von

---

1) Abweichend von Lämmels Angaben, l. c.

$F^{(i)}$, das mit der Wahrscheinlichkeit $p^{(i)}$ zu erwarten ist, für sie eine Einnahme (oder Ausgabe) $a^{(i)}$ zur Folge hat; im Falle der Einnahme soll $a^{(i)}$ positiv, im andern Falle negativ sein. Dann hat die Person eine ungewisse Summe x zu gewärtigen, die $n+1$ verschiedener Werte, eines jeden mit bestimmter Wahrscheinlichkeit, fähig ist, und man kann im Sinne von Nr. 55 von einem Durchschnittswert dieser Summe sprechen, der sich als Summe der Produkte ihrer möglichen Werte mit den zugehörigen Wahrscheinlichkeiten darstellt; wird er mit E bezeichnet, so ist

$$E = pa + p'a' + p''a'' + \cdots + p^{(n)}a^{(n)}, \tag{1}$$

wobei gleichzeitig

$$p + p' + p'' + \cdots + p^{(n)} = 1 \tag{2}$$

besteht.

Ist nur auf das Ereignis F eine Summe, ein Preis a ausgesetzt, während die übrigen Ereignisse keine Vermögensänderung zur Folge haben, so kann x nur die beiden Werte a und 0 annehmen, ersteren mit der Wahrscheinlichkeit p, letzteren mit der Wahrscheinlichkeit $q = p' + p'' + \cdots + p^{(n)}$, und es ist

$$E = pa. \tag{3}$$

Den Mittelwert E bezeichnet man im vorliegenden Falle als die *mathematische Erwartung* oder *mathematische Hoffnung* der Person gegenüber den ungewissen Summen. Hiernach ist *die auf eine einzelne Summe bezügliche mathematische Erwartung das Produkt aus der Summe und der Wahrscheinlichkeit ihrer Realisierung, und die auf eine Reihe einander ausschließender Eventualsummen bezügliche mathematische Erwartung gleich der Summe der Erwartungen, welche die einzelnen Beträge betreffen.*

Man spricht in dem letzteren Falle von einer *totalen mathematischen Erwartung* oder Hoffnung gegenüber den Einzelerwartungen, aus denen sie sich zusammensetzt.

Von mathematischer Hoffnung wird häufig auch gesprochen, wenn in dem Produkt pa der Faktor a nicht eine Geldsumme, sondern irgend eine andere vom Zufall abhängige Größe bedeutet.

114. Beziehung der mathematischen Erwartung zum wahrscheinlichsten Erfolg. Um zur Bedeutung dessen zu gelangen, was soeben als mathematische Erwartung formal definiert worden ist, stellen wir folgende Betrachtung an.

Auf das Eintreffen des Ereignisses F, dem die Wahrscheinlichkeit p zukommt, während das Nichteintreffen mit der Wahrscheinlichkeit $q = 1 - p$ zu erwarten ist, sei der Preis a ausgesetzt.

Bei einer einmaligen Realisierung der Bedingungen hat die Person,

welche sich zur Zahlung des Preises verpflichtet — der Unternehmer —, entweder a oder 0 auszufolgen.

Werden s Realisierungen vereinbart und ist s eine Zahl, welche sich im Verhältnis $p:q$ in zwei Teile zerlegen läßt, so ist die wahrscheinlichste von dem Unternehmer zu leistende Zahlung spa, weil die Kombination: sp-mal F und sq-mal Nicht-F unter allen die wahrscheinlichste ist; die durchschnittlich auf eine Realisierung entfallende Zahlung ist dann pa.

Sind sp, sq nicht ganze Zahlen, so gibt es (vgl. Nr. 67) zwischen $sp - q$ und $sp + p$ eine ganze Zahl, die in der Form $sp + \delta$ (mit $|\delta| < 1$) geschrieben werden kann und so beschaffen ist, daß $(sp + \delta)$-mal F und $(sq - \delta)$-mal Nicht-F die wahrscheinlichste Kombination bezeichnet. Die wahrscheinlichste Leistung des Unternehmers beträgt jetzt $(sp + \sigma)a$, und davon entfällt auf eine Realisierung durchschnittlich der Betrag $pa + \frac{\delta a}{s}$, der sich mit wachsendem s dem pa als Grenze nähert.

Man kann demnach den Satz aussprechen: Die wahrscheinlichste, durchschnittlich auf eine von s Realisierungen entfallende Zahlung des Unternehmers ist der mathematischen, auf eine Realisierung gerichteten Erwartung entweder genau gleich oder weicht von ihr um eine Größe der Ordnung $\frac{1}{s}$ ab.

115. Die aus einer wiederholten Realisierung resultierende mathematische Erwartung. Die Bedingungen seien dieselben wie in der vorigen Nummer. Aus s Wiederholungen können, entsprechend den möglichen Kombinationen des Eintreffens und Ausbleibens von F, $s + 1$ verschiedene Erfolge erzielt werden:

$$sa,\ (s-1)a,\ \cdots\ (s-r)a,\ \cdots\ a,\ 0;$$

ihre Wahrscheinlichkeiten sind durch die Glieder der Entwicklung von $(p+q)^s$ bestimmt; insbesondere kommt dem Erfolge $(s-r)a$ die Wahrscheinlichkeit $\binom{s}{r} p^{s-r} q^r$ zu. Mithin ist die aus der wiederholten Realisierung hervorgehende mathematische Erwartung

$$E = a \sum_{0}^{s} {}_r (s-r) \binom{s}{r} p^{s-r} q^r.$$

Um den Wert der Summe zu ermitteln, gehe man von der allgemeineren Summe

$$\sum_{0}^{s} \binom{s}{r} (tp)^{s-r} q^r = (tp+q)^s$$

aus, in welcher t eine Hilfsvariable bedeutet; durch Differentiation nach dieser entsteht

$$\sum_0^s (s-r)\,p\binom{s}{r}(tp)^{s-r-1}q^r = sp\,(tp+q)^{s-1},$$

und hieraus ergibt sich, wenn $t = 1$ gesetzt wird:

$$\sum_0^s (s-r)\binom{s}{r}p^{s-r}q^r = sp;$$

demnach ist

$$E = spa,$$

also die s-fache auf die einfache Entscheidung gerichtete mathematische Erwartung.

Da sich die auf mehrere voneinander unabhängige und einander ausschließende Ereignisse bezüglichen Erwartungen bei der Bildung der Gesamterwartung summieren, so kann auch für den allgemeinen Fall: Auf die Ereignisse $F, F', \cdots F^{(n)}$, deren Wahrscheinlichkeiten $p, p', \cdots p^{(n)}$ sich zur Einheit ergänzen, sind die Preise $a, a', \cdots a^{(n)}$ ausgesetzt — die auf die s-malige Verwirklichung der allgemeinen Bedingungen gerichtete mathematische Hoffnung angegeben werden; sie ist

$$E = s(pa + p'a' + \cdots + p^{(n)}a^{(n)}).$$

116. Beziehungen zwischen Preis und Einsatz; Gewinnteilungsregel. Eine Person — der Unternehmer — setze auf das Eintreffen eines Ereignisses F von der Wahrscheinlichkeit p einen Preis a aus und gehe diese Verpflichtung s-mal, derselben Person oder verschiedenen Personen — den Spielern — gegenüber, ein. Dann ist spa die wahrscheinlichste Zahlung, welche der Unternehmer zu leisten haben wird. Soll ihm daraus weder ein Nutzen noch ein Schaden erwachsen, so wird die gleiche Summe seitens der Spieler an ihn abzuführen sein, d. h. der auf eine Entscheidung zu leistende Einsatz ist mit pa zu bemessen.

Die auf den ausgesetzten Preis bezügliche mathematische Erwartung des Spielers bezeichnet also seinen rechtmäßigen Einsatz.

Hat eine Person die Anwartschaft, den Preis a mit der Wahrscheinlichkeit p zu gewinnen, so repräsentiert diese Anwartschaft vor der Entscheidung einen Besitz, der mit demselben Werte zu bemessen wäre, welchen die Person als Einsatz zu leisten hätte, also mit der mathematischen Erwartung pa; dieser Betrag würde auch den rechtmäßigen Kaufpreis darstellen, um welchen sie die Anwartschaft an eine andere Person abgeben könnte. Nicht aber darf die mathematische Erwartung als derjenige Betrag erklärt werden, den die betreffende Person vor der Entscheidung als *sicheren Besitz* anzusehen

hat.[1]) Der mathematischen Hoffnung kommt im Grunde genommen für den Fall einer einmaligen Realisierung keine reale Bedeutung zu: aus einer solchen kann der Person nur eine der Vermögensänderungen 0 und a, niemals aber der Vermögenszuwachs pa erwachsen. Erst wenn es sich um eine sehr große Anzahl von Realisierungen, um ein Massenspiel, handelt, erlangt pa eine reale Bedeutung, indem es dann die *durchschnittlich* auf eine Realisierung entfallende wahrscheinlichste Vermögensänderung bedeutet. Die Wahrscheinlichkeitsrechnung kann daher zur „Regelung" eines Einzelspiels nichts aussagen, ein solches bleibt immer Sache der persönlichen Entschließung, unter Umständen ein Wagnis; hingegen ist sie zur Regelung von Massenspielen, worunter auch alle auf zufällige Ereignisse gestützten ernsthaften Unternehmungen gemeint sein sollen, berufen.

Der Person, welche auf den Preis a den Einsatz pa leistet, fällt, wenn sie den Preis erzielt, ein Gewinn $g = a - pa$ zu, weil sie den gezahlten Einsatz nicht zurückerhält; zwischen Gewinn und Preis besteht somit die Beziehung

$$g = qa;$$

zwischen Gewinn und Einsatz $E = pa$ ergibt sich mithin die Relation

$$\frac{E}{g} = \frac{p}{q}.$$

Dieselbe läßt auch folgende Deutung zu: Zu dem Preise a steuert der Spieler den Einsatz E, der Unternehmer den Gewinn g bei; der Spieler erlangt den Gewinn mit der Wahrscheinlichkeit p, der Unternehmer den Einsatz mit der Wahrscheinlichkeit q. *In einem geordneten Spiele oder einer Wette verhalten sich demnach die Einzahlungen der beiden Partner wie ihre Wahrscheinlichkeiten, das Spiel oder die Wette zu gewinnen.* Nachdem man die Proportion in die Form $pg = qE$ gebracht, kann man auch sagen, daß bei einem geordneten Spiele die mathematischen Erwartungen der Partner in bezug auf die ihnen in Aussicht stehenden Gewinne einander gleich sind, oder daß die totale mathematische Hoffnung eines jeden Partners gleich Null ist; denn diese Hoffnung drückt sich für den Spieler durch $pg - qE$, für den Unternehmer durch $qE - pg$ aus.

Steht in einem Spiele, das eine Person mit der Wahrscheinlichkeit p, ihr Partner mit der Wahrscheinlichkeit $q = 1 - p$ zu gewinnen erwartet, die Summe a, und verzichten die Spieler darauf, den Zufall entscheiden zu lassen, so haben sie sich rechtmäßig in die Summe a derart zu teilen, wie sie zu ihr hätten beitragen müssen, d. h. im Ver-

---

1) Vgl. hiermit Poisson, Recherches sur la probabil., deutsch von H. Schnuse, p. 42.

hältnisse der Wahrscheinlichkeiten p, q. Hierin ist einer der ältesten Sätze der Wahrscheinlichkeitstheorie enthalten, der dem Pascalschen Teilungsproblem (Nr. 40) zugrunde liegt.

117. Beispiel LI. Die Bestimmung der mathematischen Erwartung für einen Glücksfall, in welchem mehrere verschiedene Erfolge eintreten können, erfordert nach Nr. 113 die Feststellung dieser Erfolge und die Berechnung ihrer Wahrscheinlichkeiten; gerade in der letzten Forderung liegt mitunter eine erhebliche Schwierigkeit der Aufgabe, die manchmal dadurch umgangen werden kann, daß man den Glücksfall in einen andern, ihm bezüglich der mathematischen Hoffnung äquivalenten umwandelt, der eine einfachere Berechnung zuläßt. Einige Beispiele dieser Art sollen nun vorgeführt werden. Das erste bestehe in folgendem:

Eine Urne enthält n Kugeln, die mit den Nummern $1, 2, \cdots n$ bezeichnet sind. Eine Person, welche die Kugeln nach und nach zieht, erhält jedesmal 1 Frc., wenn die Nummer der Kugel mit der Ordnungszahl des Zuges übereinstimmt. Wie groß ist ihre Hoffnung?

Statt die Wahrscheinlichkeiten zu rechnen, daß dies $1, 2, \cdots n$-mal geschehen werde, mache man sich klar, daß *vor* den Ziehungen *jede* Kugel die Wahrscheinlichkeit $p = \frac{1}{n}$ hat, ihrer Nummer entsprechend gezogen zu weden.

Hält man nämlich in der Reihe $1, 2, 3, \cdots n$ eine bestimmte Nummer fest und permutiert die andern auf alle möglichen Arten, so erhält man die Anzahl $(n-1)!$ der günstigen Fälle, in welchen die festgehaltene Nummer an ihrer richtigen Stelle steht; und da die Anzahl der möglichen Fälle $n!$ beträgt, so ist tatsächlich für jede Nummer $\frac{(n-1)!}{n!} = \frac{1}{n}$ die Wahrscheinlichkeit, an der ihrem Range entsprechenden Stelle zu erscheinen.

Mithin ist der obige Fall in bezug auf die mathematische Erwartung äquivalent dem folgenden: Es liegen n Urnen, mit den Nummern $1, 2, \cdots n$ bezeichnet, vor; in jeder derselben befinden sich n Kugeln, die ebenfalls die Nummern $1, 2, \cdots n$ tragen; man zieht aus jeder Urne eine Kugel und erhält, so oft die Nummer der Kugel mit der Nummer der Urne übereinstimmt, 1 Frc. Da nun die Übereinstimmung der Nummern bei jedem Zuge mit der Wahrscheinlichkeit $\frac{1}{n}$ zu erwarten ist, so ist die mathematische Erwartung $n \cdot \frac{1}{n} \cdot 1$ Frc. $= 1$ Frc.

Daß jedoch die beiden Glücksfälle im übrigen nicht identisch sind, geht schon aus folgender Bemerkung hervor. Den größtmöglichen Erfolg von n Frcs. zu erzielen, hat bei der ersten Modalität die Wahrscheinlichkeit $\frac{1}{n!}$, bei der zweiten die kleinere Wahrscheinlichkeit $\frac{1}{n^n}$.

118. Beispiel LII. *Aus einer Urne, welche weiße und schwarze Kugeln in solchem Mengenverhältnis enthält, daß für das Ziehen einer weißen die Wahrscheinlichkeit p, für das Ziehen einer schwarzen die Wahrscheinlichkeit $q = 1 - p$ besteht, werden n Ziehungen gemacht, wobei die gezogene Kugel jedesmal zurückgelegt wird. So oft eine weiße Kugel erscheint, der eine schwarze vorausging und nachfolgt, erhält der Spieler 1 Frc. Wie groß ist die Hoffnung?*

Die Lösung dieser komplizierten Frage wird einfach, sobald man sich die Überzeugung verschafft hat, daß *vor* Beginn der Ziehungen der beschriebene Glücksfall an jeder Stelle, die erste und letzte ausgenommen, zu erwarten ist mit der Wahrscheinlichkeit pq^2.

Die Kombination schwarz-weiß-schwarz hat, für sich betrachtet, an jeder Stelle der Ziehungsreihe die Wahrscheinlichkeit pq^2. Diese Wahrscheinlichkeit wäre nun nach und nach mit den Wahrscheinlichkeiten der verschiedenen Kombinationen zu multiplizieren, die sich in dem übrigen Teil der Ziehungsreihe zutragen können, und die Summe der Produkte gäbe die Wahrscheinlichkeit des Gewinnens an der betreffenden Stelle. Hebt man aber pq^2 aus der Summe heraus, so ist der zweite Faktor die Entwicklung von $(p + p)^{n-3}$, hat also den Wert 1.

In bezug auf die mathematische Erwartung ist das Spiel also äquivalent $n - 2$ Ziehungen aus einer Urne, welche dem Erscheinen einer weißen Kugel die Wahrscheinlichkeit pq^2 verleiht, wenn auf die weiße Kugel ein Preis von 1 Frc. ausgesetzt ist. Die hierauf bezügliche Hoffnung ist aber $(n - 2)pq^2$.[1])

Über diesen Punkt geht aber die Übereinstimmung nicht hinaus; während bei der zweiten Modalität $n - 2$ Frcs. als höchster Preis erzielbar sind, beträgt bei dem vorgelegten Spiele der höchste Preis $\frac{n-1}{2}$ oder $\frac{n}{1} - 1$, je nachdem n ungerad oder gerad ist.

119. Beispiel LIII. *Peter hat m, Paul n Taler; jeder wirft seine Taler hin, und demjenigen, bei dem Wappen öfter erscheint, fallen alle $m + n$ Taler zu. Entspricht das Spiel dem Grundsatze der Billigkeit?*

Von den beiden Zahlen sei m die größere. Es kommt darauf an, für jeden der beiden Spieler die Wahrscheinlichkeit zu bestimmen, daß er den andern in dem bezeichneten Sinne übertreffen werde.

Wirft Peter einmal Wappen, so darf bei Paul kein Taler Wappen zeigen; die zusammengesetzte Wahrscheinlichkeit hierfür ist

$$\binom{m}{1}\left(\frac{1}{2}\right)^m\left(\frac{1}{2}\right)^n.$$

1) Vgl. die abweichende Lösung J. Bertrands in seinem Calcul des probabilités, p. 52.

Wirft Peter zweimal Wappen, so darf Paul höchstens einmal Wappen treffen, und dies ist mit der Wahrscheinlichkeit

$$\binom{m}{2}\left(\frac{1}{2}\right)^m\left[\left(\frac{1}{2}\right)^n+\binom{n}{1}\left(\frac{1}{2}\right)^n\right]$$

zu erwarten.

Wirft Peter dreimal Wappen, so darf dies bei Paul höchstens zweimal eintreten; die Wahrscheinlichkeit dafür ist

$$\binom{m}{3}\left(\frac{1}{2}\right)^m\left[\left(\frac{1}{2}\right)^n+\binom{n}{1}\left(\frac{1}{2}\right)^n+\binom{n}{2}\left(\frac{1}{2}\right)^n\right].$$

Nach diesem Gesetze hat man fortzuschreiten, bis man zu dem Falle, wo Peter n-mal und Paul höchstens $n-1$-mal Wappen getroffen hat, wofür die Wahrscheinlichkeit

$$\binom{m}{n}\left(\frac{1}{2}\right)^m\left[\left(\frac{1}{2}\right)^n+\binom{n}{1}\left(\frac{1}{2}\right)^n+\cdots+\binom{n}{n-1}\left(\frac{1}{2}\right)^n\right]$$

besteht. Von da ab, d. h. wenn Peter mehr als n Wappen erzielt, gewinnt er ohne Rücksicht auf den Erfolg Pauls; demnach ist die totale Wahrscheinlichkeit für Peters Gewinn:

$$\begin{aligned} p_1 = \frac{1}{2^{m+n}}\Big[\binom{m}{1} &+ \binom{m}{2}\left\{1+\binom{n}{1}\right\}+\binom{m}{3}\left\{1+\binom{n}{1}+\binom{n}{2}\right\}+\cdots \\ &+\binom{m}{n}\left\{1+\binom{n}{1}+\cdots+\binom{n}{n-1}\right\}\Big] \\ &+\frac{1}{2^m}\left[\binom{m}{n+1}+\binom{m}{n+2}+\cdots+1\right]; \end{aligned}$$

für Paul ergibt sich durch eine ähnliche Betrachtung die Wahrscheinlichkeit

$$\begin{aligned} p_2 = \frac{1}{2^{m+n}}\Big[\binom{n}{1} &+ \binom{n}{2}\left\{1+\binom{m}{1}\right\}+\binom{n}{3}\left\{1+\binom{m}{1}+\binom{m}{2}\right\}+\cdots \\ &+\binom{n}{n}\left\{1+\binom{m}{1}+\cdots+\binom{m}{n-1}\right\}\Big]. \end{aligned}$$

Ist $\frac{p_1}{p_2}=\frac{m}{n}$, so ist das Spiel billig; ist $\frac{p_1}{p_2}>\frac{m}{n}$, so ist Peter, und im Falle $\frac{p_1}{p_2}<\frac{m}{n}$ Paul im Vorteile.

Wenn $m=3$ und $n=2$, so geben die Formeln

$$p_1=\frac{1}{2},\quad p_2=\frac{3}{16},$$

und weil $\frac{8}{3}>\frac{3}{2}$, so ist Peter im Vorteile.

Wenn $m = 4$, $n = 3$ ist, so hat man

$$p_1 = \frac{1}{2}, \quad p_2 = \frac{29}{128}$$

und da $\frac{64}{29} > \frac{4}{3}$, so ist auch jetzt Peter im Vorteile.

Wird vereinbart, daß das Spiel, wenn es das erste Mal unentschieden bleiben sollte, fortzusetzen ist bis zur Entscheidung, dann kommen die *relativen* Wahrscheinlichkeiten in Betracht, nämlich $\frac{p_1}{p_1 + p_2}$ für Peter und $\frac{p_2}{p_1 + p_2}$ für Paul. In betreff des Schlusses auf die Billigkeit des Spieles bleibt diese Abänderung ohne Einfluß; die mathematische Erwartung erfährt aber eine Änderung. Kann das Spiel auch unentschieden bleiben, in welchem Falle jeder seine Taler zurücknimmt, so ist Peters Erwartung gleich

$$p_1(m + n) + (1 - p_1 - p_2)m = p_1 n + (1 - p_2)m;$$

wird es fortgesetzt bis zur Entscheidung, so beträgt Peters Erwartung $\frac{p_1}{p_1 + p_2}(m + n)$ Taler; für den letztbehandelten Spezialfall sind diese Hoffnungen beziehungsweise $4\frac{19}{32}$ und $4\frac{76}{93}$ Taler und übertreffen beide Peters Einsatz.

120. Die Sätze von Tchebycheff. Tchebycheff hat einige Wahrscheinlichkeitssätze über die Mittelwerte dem Zufall unterworfener Größen[1]) abgeleitet, die sich neben der elementaren Begründung durch einen hohen Grad von Allgemeinheit auszeichnen und eben deshalb geeignet sind, die Lösung vieler Fragen der Wahrscheinlichkeitstheorie auf sie zurückzuführen.

Es seien $x, y, z \cdots$ irgend welche Größen, deren jede mehrere verschiedene Werte, einen jeden mit bestimmter Wahrscheinlichkeit, annehmen kann, und zwar sei x der Werte

$$x_1, x_2, \cdots x_k$$

mit den Wahrscheinlichkeiten

$$p_1, p_2, \cdots p_k$$

fähig, wobei

$$\sum_1^k p_\varkappa = 1; \tag{1}$$

ferner y der Werte

$$y_1, y_2, \cdots y_l$$

---

1) Journ. Liouv. (2) XII, 1867, p. 177.

mit den Wahrscheinlichkeiten

$$q_1, q_2, \cdots q_l,$$

wobei wieder

$$\sum_1^l q_\lambda = 1; \tag{2}$$

dann z der Werte

$$z_1, z_2, \cdots z_m$$

mit den Wahrscheinlichkeiten

$$r_1, r_2, \cdots r_m,$$

deren Summe

$$\sum_1^m r_\mu = 1, \tag{3}$$

usw.

Dann sind

$$\sum_1^k p_\varkappa x_\varkappa = a, \quad \sum_1^l q_\lambda y_\lambda = b, \quad \sum_1^m r_\mu z_\mu = c, \cdots \tag{4}$$

die auf $x, y, z, \cdots$ bezüglichen mathematischen Erwartungen oder die Mittelwerte dieser Größen,

$$\sum_1^k p_\varkappa x_\varkappa^2 = a_1, \quad \sum_1^l q_\lambda y_\lambda^2 = b_1, \quad \sum_1^m r_\mu z_\mu^2 = c_1, \cdots \tag{5}$$

die Mittelwerte ihrer Quadrate.

Daß die Realisierung der Bedingungen die Wertekombination $x_\varkappa, y_\lambda, z_\mu, \cdots$ von $x, y, z, \cdots$ und daher den Wert $x_\varkappa + y_\lambda + z_\mu + \cdots$ für die Summe $x + y + z + \cdots$ herbeiführen werde, ist mit der Wahrscheinlichkeit $p_\varkappa q_\lambda r_\mu \cdots$ zu erwarten, und dieselbe Wahrscheinlichkeit besteht auch für das Stattfinden der Differenz

$$x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots$$

zwischen der beobachteten Summe und der Summe der Mittelwerte. Demnach ist der Mittelwert des Quadrates dieser Differenz durch die vielfache Summe

$$\sum (x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots)^2 p_\varkappa q_\lambda z_\mu \cdots \qquad \begin{cases} \varkappa = 1, 2, \ldots k \\ \lambda = 1, 2, \ldots l \\ \mu = 1, 2, \ldots m \\ \cdots\cdots\cdots \end{cases}$$

ausgedrückt. Schreibt man die Differenz in der Form

$$(x_\varkappa - a) + (y_\lambda - b) + (z_\mu - c) + \cdots,$$

so gibt ihr Quadrat zweierlei Glieder: quadratische und produktförmige in den Differenzen $x_\varkappa - a$, $y_\lambda - b$, $z_\mu - c, \cdots$.

Die Summierung auf das erste quadratische Glied

$$(x_\varkappa - a)^2 = x_\varkappa^2 - 2ax_\varkappa + a^2$$

zunächst in bezug auf den Zeiger $\varkappa$ ausgeführt gibt

$$\sum_1^k (x_\varkappa^2 - 2ax_\varkappa + a^2)p_\varkappa = \sum p_\varkappa x_\varkappa^2 - 2a\sum p_\varkappa x_\varkappa + a^2\sum p_\varkappa$$
$$= a_1 - 2a^2 + a^2 = a_1 - a^2;$$

die Summierung in bezug auf die andern Zeiger ändert an diesem Werte nichts, weil vermöge (2), (3), $\cdots$

$$\sum q_\lambda r_\mu \cdots = 1 \text{ (für } \lambda = 1, 2, \cdots l;\ \mu = 1, 2 \cdots m; \cdots).$$

Ebenso ergibt das zweite quadratische Glied $b_1 - b^2$, das dritte $c_1 - c^2$, usw.

Bei dem ersten produktförmigen Gliede

$$2(x_\varkappa - a)(y_\lambda - b) = 2(x_\varkappa y_\lambda - bx_\varkappa - ay_\lambda + ab)$$

ergibt die Summierung in bezug auf den Zeiger $\varkappa$

$$2\left(y_\lambda \sum p_\varkappa x_\varkappa - b\sum p_\varkappa x_\varkappa - ay_\lambda \sum p_\varkappa + ab\sum p_\varkappa\right)$$
$$= 2(ay_\lambda - ab - ay_\lambda + ab) = 0,$$

und die weiteren Summierungen können an diesem Werte nichts ändern; in gleicher Weise verschwinden die aus den übrigen produktförmigen Gliedern entspringenden Anteile der Summe.

Somit ist

$$\sum(x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots)^2 p_\varkappa q_\lambda r_\mu \cdots$$
$$= a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots; \tag{6}$$

daraus folgt aber, daß

$$\frac{\sum(x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots)^2 p_\varkappa q_\lambda r_\mu \cdots}{\alpha^2(a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots)} = \frac{1}{\alpha^2}.$$

Läßt man in der Summe, welche den Zähler bildet, alle Glieder fort, in denen

$$\frac{(x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots)^2}{\alpha^2(a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots)} < 1 \tag{7}$$

ist, und ersetzt den linksstehenden Quotienten, so oft er größer ist als 1, durch die Einheit, so ist durch beide Prozesse die linke Seite vermindert worden, mithin das Verbleibende, d. i.

$$\sum p_\varkappa q_\lambda r_\mu \cdots < \frac{1}{\alpha^2},$$

wobei sich, dem Ausgeführten zufolge, die Summierung nur auf solche Kombinationen von $\varkappa, \lambda, \mu, \cdots$ erstreckt, für welche

$$\frac{(x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots)^2}{\alpha^2(a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots)} \geqq 1. \qquad (8)$$

Die so eingeschränkte Summe $\sum p_\varkappa q_\lambda r_\mu \cdots$ bedeutet aber die Wahrscheinlichkeit $1 - P$ für das Stattfinden eben dieser Relation (8), so daß P die Wahrscheinlichkeit für das Stattfinden von (7) vorstellt. Da nun $1 - P < \frac{1}{\alpha^2}$, so ist $P > 1 - \frac{1}{\alpha^2}$. Hiernach ergibt sich der Satz:

I. *Die Wahrscheinlichkeit P, daß die Differenz*

$$x_\varkappa + y_\lambda + z_\mu + \cdots - a - b - c - \cdots$$

zwischen den Grenzen

$$-\alpha\sqrt{a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots}$$

und

$$\alpha\sqrt{a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots},$$

also die Summe

$$x_\varkappa + y_\lambda + z_\mu + \cdots$$

zwischen den Grenzen

$$a + b + c + \cdots - \alpha\sqrt{a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots}$$

und

$$a + b + c + \cdots + \alpha\sqrt{a_1 + b_1 + c_1 + \cdots - a^2 - b^2 - c^2 - \cdots}$$

enthalten sei, ist größer als $1 - \frac{1}{\alpha^2}$.

Die Zahl α ist nur an die Bedingung gebunden, daß sie größer sein muß als 1.

Ist n die Anzahl der Größen $x, y, z, \cdots$ und setzt man

$$\alpha = \frac{\sqrt{n}}{t},$$

so ergibt sich als eine Folgerung von I unmittelbar der folgende Satz:

II. *Die Wahrscheinlichkeit P, daß das arithmetische Mittel*

$$\frac{x_\varkappa + y_\lambda + z_\mu + \cdots}{n}$$

aus den beobachteten Werten von $x, y, z \cdots$ enthalten sei zwischen den Grenzen

$$\frac{a + b + c + \cdots}{n} - \frac{1}{t}\sqrt{\frac{a_1 + b_1 + c_1 + \cdots}{n} - \frac{a^2 + b^2 + c^2 + \cdots}{n}}$$

und

$$\frac{a+b+c+\cdots}{n} + \frac{1}{t}\sqrt{\frac{a_1+b_1+c_1+\cdots}{n} - \frac{a^2+b^2+c^2+\cdots}{n}},$$

ist größer als $1 - \frac{t^2}{n}$.

Wenn die Mittelwerte $a, b, c, \cdots a_1, b_1, c_1, \cdots$ eine feste Grenze nicht überschreiten, so bleiben auch die arithmetischen Mittel $\frac{a_1+b_1+c_1+\cdots}{n}$, $\frac{a^2+b^2+c^2+\cdots}{n}$, wie groß auch n sein möge, unter einer festen Grenze, somit auch die Wurzelgröße, welche in dem Satze II vorkommt. Daraus folgt, daß man die letztgenannten Grenzen durch Wahl von t allein beliebig eng ziehen kann; und da bei beliebig großem t die Differenz $1 - \frac{t^2}{n}$ mit beständig wachsendem n der Einheit als Grenze sich nähert, so ist der folgende Satz erwiesen:

III. *Wenn die Mittelwerte von $x, y, z, \cdots$ und die ihrer Quadrate über gewisse feste Grenzen nicht hinausgehen, so nähert sich die Wahrscheinlichkeit P, daß der Unterschied zwischen dem arithmetischen Mittel von n beobachteten Werten der x, y, z, und dem arithmetischen Mittel ihrer Mittelwerte $a, b, c, \cdots$ kleiner sei als eine beliebige klein festgesetzte Größe, mit wachsendem n der Einheit.*

121. Folgerungen aus diesen Sätzen. Das Poissonsche und das Bernoullische Gesetz der großen Zahlen. Einer Person falle mit dem Eintreffen des Ereignisses F, das mit der Wahrscheinlichkeit p zu erwarten ist, der Preis A zu, während das Nichteintreffen von F ohne Folgen bleibt. Es werden s Realisierungen vorgenommen. Von den Erfolgen $x^{(1)}, x^{(2)}, \cdots x^{(s)}$, welche sich dabei einstellen, kann jeder einen der beiden Werte A und 0 mit der Wahrscheinlichkeit p, beziehungsweise $q = 1 - p$ annehmen. Demnach ist der mittlere Wert eines jeden $x^{(i)}$

$$a = pA,$$

der Mittelwert seines Quadrates

$$a_1 = pA^2.$$

Nach dem Satze I ist mit einer Wahrscheinlichkeit

$$P > 1 - \frac{1}{\alpha^2}$$

zu erwarten, daß die erzielte Summe $x^{(1)} + x^{(2)} + \cdots + x^{(s)}$ enthalten sein werde zwischen den Grenzen

$$spA - \alpha\sqrt{spA^2 - sp^2A^2} \quad \text{und} \quad spA + \alpha\sqrt{spA^2 - sp^2A^2}$$

oder zwischen

$$spA - \alpha A\sqrt{spq} \quad \text{und} \quad spA + \alpha A\sqrt{spq}; \tag{1}$$

und nach dem Satze II besteht eine Wahrscheinlichkeit

$$P > 1 - \frac{t^2}{s}$$

dafür, daß der durchschnittlich auf eine Realisierung entfallende Erfolg eingeschlossen sei zwischen die Grenzen

$$pA - \frac{A}{t}\sqrt{pq} \quad \text{und} \quad pA + \frac{A}{t}\sqrt{pq}. \tag{2}$$

Die Grenzen (1) erweitern sich bei gleichbleibender unterer Grenze von P im Verhältnis der Quadratwurzel aus s, die Grenzen (2) hingegen verengen sich unter den gleichen Umständen im Verhältnis der Quadratwurzel aus s, und man kann zu beliebig eng gewählten Grenzen (2) s so groß annehmen, daß die untere Grenze von P sich beliebig wenig von der Einheit unterscheide.

Der in Rede stehenden Person möge wieder für den Fall des Eintreffens von F der Preis A ausgezahlt werden; dagegen sei sie verpflichtet, bei Nichteintreffen von F den Betrag B zu erlegen. Von den Erfolgen $x^{(1)}, x^{(2)}, \cdots x^{(s)}$, die sich jetzt bei s Realisierungen einstellen können, ist jeder zweier Werte, A und $-B$, mit der Wahrscheinlichkeit p, respektive q fähig, und es ist der mittlere Wert eines jeden $x^{(i)}$ gleich

$$a = pA - qB,$$

der eines jeden $x^{(i)2}$ gleich

$$a_1 = pA^2 + qB^2.$$

Dem Satze I zufolge besteht eine den Betrag $1 - \frac{1}{\alpha^2}$ übersteigende Wahrscheinlichkeit P dafür, daß der erzielte Gesamterfolg zwischen den Grenzen

$$s(pA - qB) - \alpha\sqrt{s(pA^2 + qB^2) - s(pA - qB)^2}$$

und

$$s(pA - qB) + \alpha\sqrt{\div}$$

oder zwischen

$$\left.\begin{array}{l} s(pA - qB) - \alpha(A + B)\sqrt{spq} \\ s(pA - qB) + \alpha(A + B)\sqrt{spq} \end{array}\right\} \tag{3}$$

und

enthalten sein werde; und dem Satz II zufolge ist mit einer Wahrscheinlichkeit P, die größer ist als $1 - \frac{t^2}{s}$, zu erwarten, daß der im Durchschnitt auf eine Realisierung entfallende Erfolg nicht über die Grenzen

$$pA - qB - \frac{A + B}{t}\sqrt{pq} \quad \text{und} \quad pA - qB + \frac{A + B}{t}\sqrt{pq} \tag{4}$$

hinausfallen werde.

Aus dem Ansatze (3) geht folgender Sachverhalt hervor: Ist $pA - qB \neq 0$, so läßt sich, wie groß auch α sein, wie nahe also P der Einheit liegen möge, s so bestimmen, daß die untere sowohl als die obere Grenze das Vorzeichen von $pA - qB$ hat. Wenn hingegen die Beträge A, B so geregelt sind, daß $pA - qB = 0$ ist, dann kann der Gesamterfolg, wie groß auch s sein mag, ebensowohl einen Gewinn wie auch einen Verlust bedeuten.

Würde beispielsweise die Person $5\frac{1}{2}$ Frcs. erhalten, wenn beim Würfeln As fällt, und 1 Frcs. zu zahlen haben, wenn eine andere Würfelseite erscheint, und wollte man mit einer Wahrscheinlichkeit, die $1 - \frac{1}{100} = \frac{99}{100}$ übertrifft, erwarten dürfen, daß aus s-maliger Wiederholung des Spieles der Person ein Gewinn erwachse, so müßte $s > 84.500$ festgesetzt werden; denn soll

$$s\left(\frac{1}{6} \cdot 5\frac{1}{2} - \frac{5}{6} \cdot 1\right) - 10 \cdot 6\frac{1}{2}\sqrt{s\frac{1}{6} \cdot \frac{5}{6}} > 0$$

sein, so muß

$$s > 84.500$$

genommen werden.

Von seinem III. Satze hat Tchebycheff Gebrauch gemacht, um das Gesetz der großen Zahlen zu erweisen.

Es habe $x^{(1)}$ den Wert 1 oder 0, je nachdem ein Ereignis F im ersten Versuche eintrifft oder nicht, und die Wahrscheinlichkeiten hierfür seien p_1, $q_1 = 1 - p_1$; $x^{(2)}$ nehme einen der Werte 1 oder 0 an, je nachdem dasselbe Ereignis im zweiten Versuche eintrifft oder nicht, wofür die Wahrscheinlichkeiten p_2, $q_2 = 1 - p_2$ bestehen mögen, usw.; endlich bedeute $x^{(n)}$ einen der Werte 1 oder 0, je nachdem das Ereignis F im n-ten Versuche eintrifft oder nicht, wofür p_n und $q_n = 1 - p_n$ die bezüglichen Wahrscheinlichkeiten sein mögen. Dann bedeutet der beobachtete Wert von

$$\frac{x^{(1)} + x^{(2)} + \cdots + x^{(n)}}{n}$$

die relative Häufigkeit des Eintreffens von F in den n Versuchen; ferner sind die Mittelwerte von $x^{(1)}, x^{(2)}, \cdots x^{(n)}$ gleich

$$a^{(1)} = p_1 \cdot 1 + q_1 \cdot 0 = p_1, \quad a^{(2)} = p_2, \quad \cdots \quad a^{(n)} = p_n,$$

die Mittelwerte ihrer Quadrate gleich

$$a_1^{(1)} = p_1 \cdot 1^2 + q_1 \cdot 0^2 = p_1, \quad a_1^{(2)} = p_2, \quad \cdots \quad a_1^{(n)} = p_n.$$

Es besteht demnach eine Wahrscheinlichkeit $P > 1 - \frac{t^2}{n}$ dafür, daß

sich die relative Häufigkeit des Eintreffens von F zwischen den Grenzen

$$\frac{p_1+p_2+\cdots+p_n}{n} \mp \frac{1}{t}\sqrt{\frac{p_1+p_2+\cdots+p_n}{n} - \frac{p_1^2+p_2^2+\cdots+p_n^2}{n}}$$

befinden werde; und im Sinne des Satzes III nähert sich P mit wachsendem n der Einheit, wie eng man auch diese Grenzen festgesetzt haben mag. Dies aber ist der wesentliche Inhalt des Theorems von Poisson (s. Nr. 93).

Unter der Annahme, daß die Wahrscheinlichkeit von F konstant bleibt im Laufe der Versuche, daß also $p_1 = p_2 = \cdots = p_n = p$ und $q_1 = q_2 = \cdots = q_n = q = 1 - p$, ergibt sich das Bernoullische Theorem in der folgenden Fassung: Die relative Häufigkeit des Eintreffens von F in n Versuchen liegt mit einer Wahrscheinlichkeit P, die größer ist als $1 - \frac{t^2}{n}$, zwischen den Grenzen

$$p \mp \frac{1}{t}\sqrt{pq}.$$

§ 2. Das mathematische Risiko.

122. Begriff des mathematischen Risiko. Jedes auf den Zufall gegründete Unternehmen, wenn es mathematisch geregelt ist, genügt der Grundforderung, daß die totale mathematische Erwartung eines jeden Teilnehmers gleich Null ist.

Trotz dieses gemeinsamen Merkmals unterscheiden sich aber verschiedene Unternehmungen voneinander je nach dem Verhältnis der dabei in Betracht kommenden Wahrscheinlichkeiten und nach der Höhe der auf dem Spiele stehenden Summen, und zwar in bezug auf die Aussicht auf Erzielung eines Reingewinnes oder in bezug auf die Gefahr eines Verlustes.

Ein Unternehmer setze auf das Eintreffen eines Ereignisses F, dem die Wahrscheinlichkeit p zukommt, einen Preis A aus und hebe dafür von dem Spieler den rechtmäßigen Einsatz

$$E = pA \tag{1}$$

ein. Nach Abschluß dieses Vertrages steht der Spieler vor zwei Eventualitäten: entweder den Reingewinn $A - E$ zu erzielen, und die hierauf bezügliche *Reingewinnhoffnung* ist

$$R = p(A - E) = p(1 - p)A = pqA = qE; \tag{2}$$

oder den Verlust des Einsatzes E zu erleiden, und die hierauf bezügliche *Verlusterwartung* ist

$$R' = qE; \tag{3}$$

beide sind sonach dem Betrage nach gleich. In einer ähnlichen Lage

befindet sich aber auch der Unternehmer: für ihn ist R' der Ausdruck für die Gewinn-, R der Ausdruck für die Verlusterwartung.

Der Betrag R oder der ihm gleiche R' ist als Maßstab für die Gefahr, welche das Unternehmen sowohl der einen wie der andern Seite bringt, unter dem Namen *Risiko* von Tetens[1]) eingeführt worden, während Wittstein[2]) ihn als *mathematisches*, Hausdorff[3]) neuerdings als *durchschnittliches Risiko* bezeichnet.

Der Quotient $\frac{R}{E}$ drückt das Maß der Gewinnerwartung, der ihm gleiche Quotient $\frac{R'}{E}$ das Maß der Verlusterwartung ohne Rücksicht auf die Höhe der Summen aus, die auf dem Spiele stehen. Man nennt den Wert dieser Quotienten das *relative Risiko*. Da nach (3)

$$\frac{R'}{E} = q$$

ist, so wächst die relative Verlusterwartung, oder wie man sagen kann, die Gefährlichkeit des Spiels geradeso wie die Wahrscheinlichkeit, zu verlieren.

Den Größen R, R' kommt auch eine unmittelbar verständliche Bedeutung zu. Es ist R' der rechtmäßige Einsatz oder die Prämie, welche der Spieler demselben oder einem andern Unternehmer zu leisten hätte, damit dieser ihm den wirklich eingetretenen Verlust ersetze, und R die Prämie, gegen welche der Unternehmer seinerseits sich gegen einen Verlust sichern könnte.

Durch diesen neuen Vertrag ist aber der Spieler keineswegs gegen *jeglichen* Verlust gesichert[4]); vielmehr hat sich nur seine Reingewinnhoffnung und im gleichen Maße seine Verlusterwartung verändert, und zwar vermindert. Denn seine Einzahlung ist jetzt $E + R'$, seine Reingewinnhoffnung daher

$$R_1 = p(A - E - R') = q^2 E,$$

die Verlusterwartung ebenfalls

$$R_1' = q \cdot qE = q^2 E,$$

weil er im Falle des Verlustes nur die ursprüngliche Einzahlung E, nicht aber auch die Prämie $R' = qE$ zurückerhält.

---

1) Johann Nicolaus Tetens, Einleitung zur Berechnung der Leibrenten und Anwartschaften. II. Teil. Leipzig, 1786.

2) Theodor Wittstein, Das mathematische Risiko der Versicherungsgesellschaften etc. Hannover, 1885.

3) F. Hausdorff, Das Risiko bei Zufallsspielen. Leipziger Ber. 49, 1897.

4) Diese irrtümliche Auffassung könnte den Ausführungen Wittsteins l. c. entnommen werden.

Es wäre also R_1' die rechtmäßige Prämie, gegen welche sich der Spieler auch gegen diesen Verlust sichern könnte; hätte er dies getan, so wäre seine Reingewinnhoffnung nur mehr

$$R_2 = p(A - E - R' - R_1') = q^2 E$$

und ihr stünde die gleich große Verlusterwartung

$$R_2' = q \cdot q^2 E = q^3 E$$

gegenüber.

Auf diese Weise könnte der Spieler durch eine fortlaufende Kette von Versicherungen sein Risiko beliebig verkleinern, und er würde sich dabei dem interesselosen Grenzfalle nähern, daß er die Summe A einzahlt und eine gleich große Summe mit Sicherheit zurückerhält; denn seine Einzahlung betrüge an der Grenze

$$E + R' + R_1' + R_2' + \cdots = E(1 + q + q^2 + q^3 + \cdots) = \frac{E}{1-q} = \frac{E}{p} = A.$$

Man hat die Größen $R, R_1, R_2, \cdots$ oder $R', R_1', R_2', \cdots$ auch als Risiko erster, zweiter, dritter, $\cdots$ Ordnung bezeichnet.

123. Fortsetzung. Ausdehnung auf mehrere Preise. Auf eine Reihe einander ausschließender Ereignisse $F_1, F_2, \cdots F_n$, deren Wahrscheinlichkeiten $p_1, p_2, \cdots p_n$ demnach der Bedingung

$$p_1 + p_2 + \cdots + p_n = 1$$

entsprechen, seien Preise $A_1, A_2, \cdots A_n$ ausgesetzt; der für diese Gewinnaussicht von einem Spieler zu leistende Einsatz ist

$$E = p_1 A_1 + p_2 A_2 + \cdots + p_n A_n.$$

Nach Erlag desselben ist seine Reingewinnhoffnung

$$R = \sum p_g (A_g - E),$$

wenn unter A_g alle jenen Preise verstanden werden, die größer sind als der Einsatz E; und seine Verlusterwartung ist

$$R' = \sum p_k (E - A_k),$$

wenn unter A_k alle jene Preise verstanden werden, die unter dem Einsatz liegen. Daß beide einander gleich sind, folgt aus

$$R - R' = \sum_1^n p_i (A_i - E) = \sum_1^n p_i A_i - E \sum_1^n p_i = E - E = 0;$$

daher ist auch

$$R = R' = \frac{1}{2} \sum_1^n p_i \mid A_i - E \mid. \tag{2}$$

In gleicher Weise ergibt sich das Risiko zweiter Ordnung:

$$R_1 = R_1' = \frac{1}{2} \sum_1^n p_i \mid A_i - E - R' \mid, \tag{3}$$

das der dritten Ordnung:

$$R_2 = R_2' = \frac{1}{2} \sum_1^n p_i \mid A_i - E - R - R_1' \mid, \tag{4}$$

usw. Die Summe $E + R' + R_1' + R_2' + \cdots$ nähert sich mit wachsendem Zeiger dem größten Preis als Grenze.

Bei demselben Einsatze E kann das Risiko R sehr verschieden ausfallen je nach der Verteilung der Preise und der Wahrscheinlichkeiten, sie zu erlangen.

Als Beispiel wählen wir denselben Fall, an welchem schon Tetens[1]) den Begriff des Risiko erklärt hat.

Auf jede Seite eines Würfels sei ein Preis von so viel Francs ausgesetzt, als sie Punkte trägt. Der vom Spieler hierfür zu bezahlende Einsatz ist

$$E = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 3\frac{1}{2} \text{ Frcs.},$$

das Risiko des Spiels beträgt

$$R = \frac{1}{6}\left(\frac{1}{2} + 1\frac{1}{2} + 2\frac{1}{2}\right) = \frac{3}{4} \text{ Frcs.},$$

das relative Risiko ist $\frac{3}{4} : 3\frac{1}{2} = \frac{3}{14}$, so daß der Spieler $\frac{3}{14}$ oder $21\frac{3}{7}\,\%$ des Einsatzes als Prämie zu zahlen hätte, um sich gegen den Verlust des Einsatzes zu sichern.

Zahlt er diese Prämie, so ist seine gesamte Einzahlung $4\frac{1}{4}$ Frcs. und wird nur mehr von den Preisen 5 und 6 Frcs. übertroffen; sein Risiko ist nur mehr

$$R_1 = \frac{1}{6}\left(\frac{3}{4} = 1\frac{3}{4}\right) = \frac{5}{12} \text{ Frcs.},$$

das relative Risiko $\frac{5}{12} : 4\frac{1}{3} = \frac{5}{52}$; er hätte also $\frac{5}{52}$ oder $9\frac{8}{13}\,\%$ des erhöhten Einsatzes als Prämie zu zahlen, um sich gegen den Verlust desselben zu versichern.

Hierdurch ginge aber sein Risiko, da die Einzahlung nun $4\frac{2}{3}$ Frcs. beträgt, über in

---

1) l. c.

$$R_2 = \frac{1}{6}\left(\frac{1}{3} + 1\frac{1}{3}\right) = \frac{5}{18} \text{ Frcs.},$$

das relative Risiko betrüge $\frac{5}{18} : 4\frac{2}{3} = \frac{5}{84}$, und er hätte $\frac{5}{84}$ oder $5\frac{20}{21}\,\%$ der Einzahlung an Prämie zu erlegen, um sich gegen den Verlust jener Einzahlung sicherzustellen usw.

Das Risiko nimmt also beständig ab, die zu leistenden Einzahlungen nähern sich aber wachsend dem höchsten Preise von 6 Frcs. als Grenze.

Wäre auf die Würfelseiten 4, 5, 6 je ein Preis von 7 Frcs., auf die andern kein Preis ausgesetzt worden, so wäre vom Spieler derselbe Einsatz von $3\frac{1}{2}$ Frcs. zu leisten; sein Risiko betrüge aber jetzt

$$R = \frac{1}{6}\left(3\frac{1}{2} + 3\frac{1}{2} + 3\frac{1}{2}\right) = 1\frac{3}{4} \text{ Frcs.},$$

das relative $1\frac{3}{4} : 3\frac{1}{2} = \frac{1}{2}$ oder $50\,\%$ des Einsatzes.

Und hätte der Unternehmer lediglich auf die Würfelseite 6 den Preis von 21 Frcs. ausgesetzt, so wäre bei dem gleichen Einsatze des Spielers, d. i. $3\frac{1}{2}$ Frcs., sein Risiko nunmehr

$$R = \frac{1}{6} \cdot 17\frac{1}{2} = 2\frac{11}{12} \text{ Frcs.},$$

er hätte somit $2\frac{11}{12} : 3\frac{1}{2} = \frac{5}{6}$ oder $83\frac{1}{3}\,\%$ des Einsatzes an Prämie zu zahlen, um sich gegen dessen Verlust zu schützen.

Es ist sonach unter den drei Formen des Spiels die erste die mindest, die dritte die am meisten gefährliche.

124. Das Risiko bei einer großen Anzahl voneinander unabhängiger, gleichartiger Einzelfälle. Der Unternehmer setze auf das mit der Wahrscheinlichkeit p zu erwartende Eintreffen des Ereignisses F den Preis A aus und hebe dafür von dem Spieler den rechtmäßigen Einsatz $E = pA$ ein. Wie groß ist sein Risiko, wenn er eine sehr große Anzahl s solcher Verträge, entweder mit demselben oder mit verschiedenen Spielern, abschließt?[1])

Die Wahrscheinlichkeit, daß das Ereignis F in den s Fällen

1) Man kann sich den Hergang etwa durch folgendes Schema verbildlichen: Es liegen s gleichmäßig mit weißen und schwarzen Kugeln gefüllte Urnen vor, deren jede dem Ziehen einer weißen Kugel die Wahrscheinlichkeit p verleiht; an jede dieser Urnen tritt eine Person heran und macht einen Zug; für jede dabei erschienene weiße Kugel zahlt der Unternehmer den Preis.

$(sp + l)$-mal eintreffe und $(sq - l)$-mal ausbleibe, ist näherungsweise ausgedrückt durch (vgl. Nr. 75)

$$\frac{1}{\sqrt{2\pi spq}} e^{-\frac{l^2}{2spq}},$$

wenn $q = 1 - p$ ist. Die von dem Unternehmer zu leistende Zahlung ist dann $(sp + l)A$, während seine Einnahme $sE = spA$ beträgt; er erleidet somit einen Verlust vom Betrage lA. Sein Risiko würde also erhalten, wenn man die Produkte

$$\frac{1}{\sqrt{2\pi spq}} lAe^{-\frac{l^2}{2spq}}$$

für alle zulässigen positiven l bildete und summierte. Diese Summe kann aber (vgl. Nr. 80, Schluß) mit hinreichender Genauigkeit durch das über *alle* positiven Werte von l ausgedehnte Integral der eben angeschriebenen Funktion ersetzt werden, so daß man hat:

$$R' = \frac{A}{\sqrt{2\pi spq}} \int_0^\infty le^{-\frac{l^2}{2spq}} dl;$$

zwecks Ausführung der Integration führe man mittels der Substitution $\frac{l}{\sqrt{2spq}} = t$ die neue Variable t ein, und man findet so

$$R' = A\sqrt{\frac{2spq}{\pi}} \int_0^\infty te^{-t^2} dt = A\sqrt{\frac{spq}{2\pi}}. \tag{1}$$

Daraus folgt das relative Risiko

$$\frac{R'}{spA} = \sqrt{\frac{1}{2\pi s}\frac{q}{p}} = 0{,}39894\sqrt{\frac{q}{sp}}. \tag{2}$$

Hiernach wächst das absolute Risiko R' proportional mit der Höhe des Preises und proportional mit der Quadratwurzel aus der Anzahl der abgeschlossenen Unternehmungen, hängt aber überdies noch von dem Größenverhältnis der Wahrscheinlichkeiten ab; es wird in letzterer Hinsicht am größten für $p = q = \frac{1}{2}$.

Das relative Risiko aber nimmt im Verhältnis der Quadratwurzel aus der Anzahl der Unternehmungen ab und kann daher durch Vergrößerung dieser Anzahl beliebig klein gemacht werden. Darin zeigt sich der wesentliche Unterschied zwischen einem einzelnen Glücksfalle und einem aus einer großen Anzahl gleichartiger Fälle bestehenden Unternehmen.

Zur Erläuterung diene folgendes Beispiel. Der Unternehmer setze auf eine bestimmte Würfelseite den Preis von 6 Frcs. aus und hebe dafür den Einsatz von 1 Frc. ein. Für einen einzelnen Fall beträgt das absolute Risiko $\frac{5}{6}$ Frc. oder $83\frac{1}{3}\%$ des Einsatzes; um es zu decken, müßte der Unternehmer von dem Partner neben dem Einsatz von 1 Frc. noch eine Prämie von $\frac{5}{6}$ Frc. einheben. Schließt der Unternehmer 500 solcher Fälle ab, so ist sein absolutes Risiko 19,94 Frcs., das relative aber sinkt auf 0,039894 oder etwas weniger als 4% des gesamten Einsatzes, so daß er, um es zu decken, von jedem Partner nur eine Zuschlagsprämie von $\frac{4}{100}$ Frc. einzuheben brauchte.

Es soll nun noch die Frage nach der Wahrscheinlichkeit erledigt werden, daß die von dem Unternehmer zu leistende Auszahlung sich von dem wahrscheinlichsten Betrage spA, der zugleich die Summe der von den Partnern gezahlten Einlagen darstellt, nicht mehr als um das k-fache Risiko nach auf- oder abwärts entfernen werde.

Dem Bernoullischen Theorem zufolge kann die Wahrscheinlichkeit dafür, daß sich die Wiederholungszahl von F zwischen die Grenzen $sp - l$ und $sp + l$ stellen werde, nahe genug durch

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2spq}}} e^{-t^2} dt$$

ausgedrückt werden. Die Zahlung des Unternehmers fällt dann zwischen die Grenzen $spA - lA$ und $spA + lA$. Soll nun

$$lA = kR' = kA\sqrt{\frac{spq}{2\pi}}$$

sein, so muß für l die nächste über

$$k\sqrt{\frac{spq}{2\pi}}$$

liegende ganze Zahl genommen werden; nimmt man für l diesen letzten Ausdruck selbst, so wird

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{k}{2\sqrt{\pi}}} e^{-t^2} dt = \Phi\left(\frac{k}{2\sqrt{\pi}}\right), \tag{3}$$

und dies ist die verlangte Wahrscheinlichkeit. Ihr halber Betrag drückt ebensowohl die Wahrscheinlichkeit eines das k-fache Risiko nicht übersteigenden Gewinnes wie die eines gleichgearteten Verlustes aus.

Da $\frac{1}{2}$ die Wahrscheinlichkeit eines Verlustes überhaupt ist, so bedeutet P auch die Wahrscheinlichkeit, daß, wofern ein Verlust eintritt, er unter dem k-fachen Risiko verbleibt.

Mittels der Tafel I findet man, daß für

$k =$	1	2	3	4	5	6
$P =$	0,31006	0,57498	0,76863	0,88945	0,95392	0,98332.

125. Das sogenannte mittlere Risiko. Wenn man jede Differenz zwischen einem Preise A_i und dem Einsatz E des Spielers als eine *Abweichung* bezeichnet, so stellt sich nach den Ausführungen in Nr. 123 das Risiko als der halbe Mittelwert der absoluten Beträge aller Abweichungen; dies ist der Inhalt der dort abgeleiteten Formel (2):

$$R = \frac{1}{2} \sum_1^n p_i \mid A_i - E \mid;$$

dagegen ist — bei einem geordneten Unternehmen — der mittlere Wert der Abweichung selbst gleich Null, d. h.

$$\sum_1^n p_i (A_i - E) = 0,$$

wie ebenfalls dort gezeigt ist.

Man kann nun daran denken, zur Charakterisierung eines vom Zufall abhängigen Unternehmens einen andern Mittelwert der Abweichungen zu verwenden; als ein solcher bietet sich schon aus praktischen Erwägungen der folgende dar:

$$M = \sqrt{\sum_1^n p_i (A_i - E)^2}; \tag{1}$$

er hebt den Zeichenunterschied der Abweichungen in einfacher Weise auf und läßt große Abweichungen stärker hervortreten, als wenn deren erste Potenzen in Rechnung gezogen werden, was eine strengere Beurteilung der Verlustgefahr erwarten läßt. Nach einer in einem andern Gebiete der Wahrscheinlichkeitsrechnung (s. Fehlertherorie) üblichen Nomenklatur wäre M als die *mittlere Abweichung* zu bezeichnen; es ist dafür auch der Name *mittleres Risiko* vorgeschlagen worden.[1])

Wenn es sich um einen einzelnen Glücksfall handelt, so leistet die Größe M nicht genau dasselbe wie das Risiko; setbstverständlich kommt ihr auch nicht dessen materielle Bedeutung zu. Angenommen,

1) S. die bereits zitierte Arbeit F. Hausdorffs. Dem Wesen nach kommt dieser Begriff schon in C. Bremikers Schrift: Das Risiko bei Lebensversicherungen (Berlin 1859) vor.

auf das mit der Wahrscheinlichkeit p zu erwartende Eintreffen von F sei der Preis A ausgesetzt, während im Falle des Nichteintreffens der Spieler seinen Einsatz $E = pA$ an den Unternehmer verliert, so ist das Risiko

$$R = qE,$$

das mittlere Risiko

$$M = \sqrt{p(A-E)^2 + qE^2} = E\sqrt{\frac{q}{p}};$$

unter allen Umständen ist also $M > R$. Wenn ferner bei zwei Glücksfällen mit demselben Einsatz die Risiken R', R'' in der Beziehung $R' > R''$ zueinander stehen, so findet auch die Relation $M' > M''$ statt. Aber das Verhältnis $\frac{M'}{M''}$ kann erheblich verschieden sein von $\frac{R'}{R''}$.

Anders gestaltet sich die Sachlage, wenn es sich um eine große Anzahl gleichartiger und voneinander unabhängiger Unternehmungen handelt; dann steht, wie gezeigt werden wird, das *mittlere* Gesamtrisiko (des Unternehmers oder der Vereinigung aller Spieler) zu dem *eigentlichen* Gesamtrisiko in einem konstanten, von den Wahrscheinlichkeiten unabhängigen Verhältnis, so daß in einem solchen Falle die eine Größe ebenso geeignet ist wie die andere, die Verlustgefahr oder Gewinnaussicht des Unternehmens zu kennzeichnen.

Schließt nämlich der Unternehmer s Verträge von der oben beschriebenen Art ab, wobei s eine sehr große Zahl sein soll, so ist die Wahrscheinlichkeit, daß seine Leistung $spA - lA$ *oder* $spA + lA$ betragen, die Abweichung also, vom Vorzeichen abgesehen, lA sein werde, näherungsweise durch

$$\frac{2}{\sqrt{2\pi spq}}\, e^{-\frac{l^2}{2spq}}$$

ausgedrückt (s. Nr. 75). Man erhielte also M^2 durch Multiplikation dieses Ausdrucks mit $(lA)^2$ und Summierung des Produktes über alle zulässigen Werte von l. Statt dessen kann mit zureichender Genauigkeit das Integral

$$\frac{2}{\sqrt{2\pi spq}} \int_0^\infty (lA)^2 e^{-\frac{l^2}{2spq}}\, dl$$

genommen werden; führt man hierin die neue Variable t mittels der Substitution

$$\frac{l}{\sqrt{2spq}} = t$$

ein, so ergibt sich nach Vollziehung der Integration

$$M^2 = spqA^2,$$

woraus dann

$$M = A\sqrt{spq} \tag{2}$$

folgt. Da nun gemäß der Formel (1) der vorigen Nummer das durchschnittliche Risiko für denselben Fall

$$R = A\sqrt{\frac{spq}{2\pi}}$$

beträgt, so ist tatsächlich das Verhältnis $\frac{R}{M}$ konstant, und zwar ist $\frac{1}{\sqrt{2\pi}} = 0{,}39894$ sein Wert, so daß

$$R = M\sqrt{\frac{1}{2\pi}} = 0{,}39894\, M \tag{3}$$

ist.

Das Rechnen mit der mittleren Abweichung bietet den Vorteil, daß man die mittlere Abweichung, welche aus einer Reihe verschiedenartiger Unternehmungen resultiert, die voneinander unabhängig sind, leicht berechnen kann, sobald die den einzelnen Unternehmungen entsprechenden mittleren Abweichungen bestimmt sind. Nach den einleitenden Entwicklungen zu den Tchebycheffschen Sätzen, Nr. 120, ist nämlich der Mittelwert des Quadrates einer Summe von Größen, deren Mittelwerte selbst einzeln Null sind, gleich der Summe der mittleren Quadrate der einzelnen Größen, wie sich dies aus der dortigen Gleichung (6) ergibt, wenn man darin $a = 0$, $b = 0$, $c = 0, \cdots$ setzt.

Hat demnach jemand mehrere voneinander unabhängige Verträge abgeschlossen, aus welchen die mittleren Abweichungen $M_1, M_2, M_3, \cdots$ hervorgehen, so ist die totale mittlere Abweichung oder das mittlere Gesamtrisiko M durch die Gleichung

$$M = \sqrt{M_1^2 + M_2^2 + M_3^2 + \cdots} \tag{4}$$

gegeben. Über die wirklich beobachtete Totalabweichung aber läßt sich nach dem ersten Satze von Tchebycheff das folgende Wahrscheinlichkeitsurteil aufstellen:

Die Wahrscheinlichkeit, daß die aus mehreren Unternehmungen mit den mittleren Abweichungen M_1, M_2, M_3, $\cdots$ hervorgehende Gesamtabweichuny zwischen die Grenzen

$$-\alpha\sqrt{M_1^2 + M_2^2 + M_3^2 + \cdots} \text{ und } \alpha\sqrt{M_1^2 + M_2^2 + M_3^2 + \cdots}$$

falle, ist größer als $1 - \frac{1}{\alpha^2}\cdot$

Besteht jede der Unternehmungen selbst wieder aus einer großen Anzahl gleichartiger Fälle, so können die M_1, M_2, M_3, $\cdots$ nach der Formel (2) gerechnet werden. Zwischen M und dem Gesamtrisiko R besteht auch dann die durch (3) ausgedrückte Beziehung.

Um dies zu erweisen, denken wir uns zwei Unternehmungen mit s_1, beziehungsweise s_2 Verträgen, mit den Preisen A_1, A_2 und den Wahrscheinlichkeiten p_1, p_2, diese zu erlangen. Die Wahrscheinlichkeit, daß sich die Abweichung $l_1 A_1 + l_2 A_2$ einstelle, ist bestimmt durch das Produkt

$$\frac{1}{\sqrt{2\pi s_1 p_1 q_1}}\, e^{-\frac{l_1^2}{2 s_1 p_1 q_1}} \cdot \frac{1}{\sqrt{2\pi s_2 p_2 q_2}}\, e^{-\frac{l_2^2}{2 s_2 p_2 q_2}};$$

dies Produkt, mit $l_1 A_1 + l_2 A_2$ multipliziert und über alle Werte von l_1, l_2 integriert, für welche $l_1 A_1 + l_2 A_2$ positiv ist, gibt das Risiko R.

Setzt man zur Abkürzung

$$\frac{1}{\sqrt{2 s_1 p_1 q_1}} = h_1, \quad \frac{1}{\sqrt{2 s_2 p_2 q_2}} = h_2,$$

so wird hiernach

$$R = \frac{h_1 h_2}{\pi} \int\int e^{-h_1^2 l_1^2 - h_2^2 l_2^2} (l_1 A_1 - l_2 A_1) dl_1\, dl_2$$

$$= \frac{h_1 h_2}{\pi} \left\{ A_1 \int_{-\infty}^{\infty} e^{-h_2^2 l_2^2} dl_2 \int_{-\frac{l_2 A_2}{A_1}}^{\infty} e^{-h_1^2 l_1^2} l_1 dl_1 + A_2 \int_{-\infty}^{\infty} e^{-h_1^2 l_1^2} dl_1 \int_{-\frac{l_1 A_1}{A_2}}^{\infty} e^{-h_2^2 l_2^2} l_2 dl_2 \right\}$$

$$= \frac{h_1 h_2}{\pi} \left\{ \frac{A_1^2}{2 h_1^2 \sqrt{h_1^2 A_2^2 + h_2^2 A_1^2}} \sqrt{\pi} + \frac{A_2^2}{2 h_2^2 \sqrt{h_1^2 A_2^2 + h_2^2 A_1^2}} \sqrt{\pi} \right\}$$

$$= \frac{1}{2\sqrt{\pi}} \sqrt{\frac{A_1^2}{h_1^2} + \frac{A_2^2}{h_2^2}};$$

ersetzt man h_1, h_2 wieder durch ihre Werte und nimmt Rücksicht auf Formel (2), so wird

$$R = \frac{1}{\sqrt{2\pi}} \sqrt{M_1^2 + M_2^2},$$

d. i. nach (4):

$$R = \frac{1}{\sqrt{2\pi}} M = 0{,}39894\, M, \qquad (3^*)$$

wie behauptet worden.

Sind also die Risiken der Einzelunternehmungen, deren jede aus einer großen Zahl gleichartiger Einzelfälle besteht, bekannt, so setzt sich das Gesamtrisiko nach demselben Gesetz zusammen, wie es in (4) ausgesprochen ist, daß nämlich

$$R = \sqrt{R_1^2 + R_2^2 + R_3^2 + \cdots} \qquad (4)$$

ist.

Es mag nun noch die Frage nach der Wahrscheinlichkeit beantwortet werden, daß die aus dem Unternehmen, welches durch die Gleichung (2) gekennzeichnet ist, hervorgehende Abweichung das k-fache mittlere Risiko M nicht übersteige, mit andern Worten, daß die von dem Unternehmer zu leistende Zahlung von der wahrscheinlichsten spA dem Betrage nach nicht mehr als um kM abweiche.

Nach dem Bernoullischen Theorem fällt die Wiederholungszahl von F mit der Wahrscheinlichkeit

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2spq}}} e^{-t^2} dt$$

zwischen die Grenzen $sp - l$ und $sp + l$, die zu leistende Zahlung also zwischen $spA - lA$ und $spA + lA$; soll nun

$$lA = kM = kA\sqrt{spq}$$

sein, so ist für l der Wert $k\sqrt{spq}$ zu setzen; mithin ist

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{k}{\sqrt{2}}} e^{-t^2} dt = \Phi\left(\frac{k}{\sqrt{2}}\right) \tag{5}$$

die verlangte Wahrscheinlichkeit.

Mit Hilfe der Tafel I findet man, daß die Werte

$k =$	1	2	3	4
$P =$	0,68267	0,95449	0,99730	0,99994

zusammengehören. Man hat also mit der Wahrscheinlichkeit $\frac{9973}{10000}$ zu erwarten, daß die Abweichung nicht über $3M$ betragen werde; mit der gleichen Wahrscheinlichkeit ist zu erwarten, daß, *wenn* ein Verlust eintritt, er diese Grenze nicht überschreitet.

Die Frage, welchen Fonds der Unternehmer bereit halten soll, um aus ihm eingetretene Verluste zu decken, läßt eine apodiktische Beantwortung nicht zu; absolute Sicherheit bestände nur in dem Falle, wenn jener Fonds die Höhe des größtmöglichen Verlustes hätte. Sobald er niedriger ist, besteht nur ein gewisses Maß von Wahrscheinlichkeit dafür, daß er ausreichen werde. Mit welchem Maß von Wahrscheinlichkeit man sich begnügen soll, kann aber niemals Gegenstand einer Rechnung sein, hängt vielmehr von dem subjektiven Ermessen ab. Durch Wittsteins Darstellung vom mathematischen Risiko zieht sich der Gedanke, als ob dieses das einzig richtige Maß des Sicherheitsfonds wäre, und als ob es zur Deckung *aller* Verluste ausreichte. Aus der Zusammenstellung am Schlusse der vorigen Nummer geht

indessen hervor, daß, *wenn* ein Verlust eintritt, er nur mit der Wahrscheinlichkeit $\frac{31}{100}$ unter dem Risiko verbleibt, und daß man den Fonds auf das doppelte Risiko erhöhen müßte, um mit der Wahrscheinlichkeit $\frac{575}{1000}$ erwarten zu dürfen, er werde ausreichen. Um aber solche Wahrscheinlichkeitsschlüsse über die Höhe des Fonds oder der einzuhebenden Zuschlagsprämie machen zu können, kann man sich ebenso gut auf das mittlere Risiko stützen und die obenstehende Ziffernreihe benützen.[1])

126. Beispiel LIV. *Ein Unternehmer habe folgende Verträge abgeschlossen:*

300 Vertr. auf den Prais von 2000 Frcs., zahlbar mit der Wahrsch. $\frac{1}{50}$, gegen den Einsatz von 40 Frcs.,

200 Vertr. auf den Preis von 3000 Frcs., zahlbar mit der Wahrsch. $\frac{1}{100}$, gegen den Einsatz von 30 Frcs.,

400 Vertr. auf den Preis von 4000 Frcs., zahlbar mit der Wahrsch. $\frac{1}{200}$, gegen den Einsatz von 20 Frcs.,

600 Vertr. auf den Preis von 5000 Frcs., zahlbar mit der Wahrsch. $\frac{1}{500}$, gegen den Einsatz von 10 Frcs.,

Es ist sein mittleres und sein durchschnittliches Risiko zu bestimmen.

Gemäß der Formel (2) der vorigen Nummer sind die *Quadrate* der mittleren Risiken in den vier Gruppen:

$$4000000 \cdot 300 \cdot \frac{1}{50} \cdot \frac{49}{50} = 23520000$$

$$9000000 \cdot 200 \cdot \frac{1}{100} \cdot \frac{99}{100} = 17820000$$

$$16000000 \cdot 400 \cdot \frac{1}{200} \cdot \frac{199}{200} = 31840000$$

$$25000000 \cdot 600 \cdot \frac{1}{500} \cdot \frac{499}{500} = \underline{29940000}$$

$$103120000;$$

hieraus berechnet sich nach der Formel (4) das mittlere Gesamtrisiko

$$M = \sqrt{103120000} = 10155 \text{ Frcs.}$$

und daraus nach Formel (3*) das durchschnittliche Gesamtrisiko

$$R = 0{,}39894 \cdot 10155 = 4051 \text{ Frcs.}$$

---

1) Vgl. hierzu Hausdorff, l. c., p. 515 ff.

Da nun die gesamte Einnahme des Unternehmers

$$300 \cdot 40 + 200 \cdot 30 + 400 \cdot 20 + 600 \cdot 10 = 32000 \text{ Frcs.}$$

beträgt, so macht das mittlere Risiko beiläufig $31\frac{3}{4}$, das durchschnittliche $12\frac{7}{10}$ % der Einnahme aus. Wollte also der Unternehmer einen dem mittleren Risiko gleichkommenden Sicherheitsfonds gründen, — und er hätte dann die Wahrscheinlichkeit 0,68267, daß er damit ausreichen werde — so müßte er bei gleichmäßiger Aufteilung auf die Teilnehmer deren Einsätze um je $31\frac{3}{4}$ % erhöhen.

Bei zehnfacher Anzahl der Verträge in jeder Gruppe würde das mittlere (und das durchschnittliche) Risiko nur $\sqrt{10}$-mal, die Einzahlung aber 10-mal größer; der Prozentsatz verminderte sich dadurch im Verhältnisse $\sqrt{10} : 10$ und betrüge bezüglich des mittleren Risiko nur noch 9,6, bezüglich des durchschnittlichen 3,9 %.

127. Eine andere Auffassung des Risikoproblems. In einer Abhandlung[1]), welche das Risikoproblem vom Standpunkte des Versicherungswesens behandelt, wo ihm tatsächlich die größte Bedeutung zukommt, übt Küttner an dem bisher üblichen, dem Wesen nach auf Tetens zurückführenden Risikobegriff Kritik, spricht ihm die Eignung zur Lösung der Hauptfrage nach der Höhe des zur Deckung eventueller Verluste erforderlichen Fonds ab und schlägt einen anderen Weg vor, der nach seiner Anschauung der Natur der Sache besser entspricht.

Bevor auf den wesentlichen Inhalt seiner Theorie eingegangen wird, sei bemerkt, daß die Kritik sich zum Teil gegen Auffassungen wendet, die heute nicht mehr bestehen. Wenn früher bei einzelnen Autoren die Meinung bestand — im Vorstehenden ist beispielsweise auf Wittstein hingewiesen worden —, das durchschnittliche Risiko R sei jene Summe, durch deren Vorhandensein der Unternehmer gegen *alle* aus den Abweichungen des wirklichen von dem wahrscheinlichen Verlauf möglicherweise entstehenden Verluste gesichert sei, so hat die neuere Literatur mit dieser Auffassung gebrochen und sieht in dem durchschnittlichen und mittleren Risiko zunächst nur einen Maßstab, nach welchem verschiedene Unternehmungen auf ihre Verlustgefahr miteinander verglichen werden können, nicht aber die unmittelbare Größe des zur Sicherung gegen diese Gefahr notwendigen Fonds; eine von willkürlichen Festsetzungen freie Bestimmung dieses Fonds erweist sich, weil Wahrscheinlichkeitsurteile im Spiele sind, als unmöglich, und auch Küttner ist zu solchen Festsetzungen genötigt. Was

---

1) Das Risiko der Lebensversicherungs-Anstalten und Unterstützungskassen. Veröffentlichungen des Deutschen Vereins für Versicher.-Wissensch. Berlin 1906.

die Bekämpfung des Einzelrisiko anlangt, so kann ihr so weit zugestimmt werden, als dem Einzelrisiko eine praktische Bedeutung ebensowenig zugesprochen werden kann wie der einzelnen mathematischen Hoffnung und wie der Wahrscheinlichkeit in bezug auf den einzelnen Realisierungsfall; dagegen kann gegen das Einzelrisiko, wenn es als Mittel zur Vergleichung verschiedener Zufallsunternehmungen und als Element zur Bildung des Gesamtrisikos verwendet wird, kein Einwand erhoben werden.

Küttner bringt die Vorstellung des Risikos eines Zufallsunternehmens mit dem größten Verluste in Verbindung, der dabei eintreten kann, und sein Ziel ist die unmittelbare Bestimmung der zur Deckung des eventuell eintretenden Verlustes erforderlichen *Summe*, die er in Gegensatz stellt zur *mathematischen Hoffnung*, auf der sich der übliche Risikobegriff aufbaut. Seine Definition ist aber nicht so geartet, um darauf ohne weiteres eine Analyse gründen zu können; er versteht unter *Risiko überhaupt die Möglichkeit, daß das Massenspiel*[1]) *einen Verlust erleidet durch den nicht rechnungsmäßigen Ablauf der maßgebenden Ereignisse*, und unter *Größe des Risiko die Differenz zwischen der möglichen und der rechnungsmäßigen Ausgabe für das ganze Massenspiel.*

Die Durchführung dieser Gedanken besteht in den wesentlichen Zügen in folgendem.

Ein Unternehmer schließe mit einer großen Anzahl s von Spielern $A_i\,(i=1,\,2,\cdots s)$ Verträge ab, die sich auf eine Reihe von einander ausschließenden Möglichkeiten beziehen; die diesen Möglichkeiten entsprechenden Wahrscheinlichkeiten $p_{ik}\,(k=1,\,2,\cdots\omega)$ ergänzen sich also zur Einheit; der von dem einzelnen Spieler eingehobene Einsatz E_i entspreche seiner mathematischen Hoffnung; der Unterschied zwischen der Auszahlung im Falle k und dem Einsatze — er heiße S_{ik} — wird für den Unternehmer einen Verlust bedeuten, wenn er positiv, einen Gewinn, wenn er negativ ist; zugleich besteht bei der Anordnung des Spieles für jedes i die Relation

$$\sum_1^{\omega}{}^k p_{ik} S_{ik} = 0. \tag{1}$$

Entwickelt man das Produkt[2])

$$T = \prod_{i=1}^{s} \left(\sum_1^{\omega}{}^k p_{ik} e^{S_{ik}\tau\sqrt{-1}} \right)$$

1) Wir benützen diesen Ausdruck zur Bezeichnung eines jeden vom Zufall abhängigen Massenunternehmens.

2) Man vergleiche hierzu die auf denselben Gedanken beruhende Analyse in Nr. 93.

nach den Potenzen von e, so bedeutet der Koeffizient von $e^{x\tau\sqrt{-1}}$ nach seiner Zusammensetzung aus den p_{ik} die Wahrscheinlichkeit P_x, daß die Differenz zwischen der Summe aller Auszahlungen und der Summe aller Einsätze gleich x sei; multipliziert man also T mit $e^{-x\tau\sqrt{-1}}$ und integriert das Produkt in bezug auf τ zwischen den Grenzen $-\pi$ und π, so ergibt sich

$$\int_{-\pi}^{\pi} T d\tau = 2\pi P_x.$$

Durch Entwicklung der Exponentialgröße in dem allgemeinen Faktor von T bis auf Glieder zweiter Ordnung[1]) in τ geht dieser Faktor wegen (1) über in

$$1 - \frac{\tau^2}{2} \sum p_{ik} S_{ik}^2,$$

und entwickelt man weiter lT bis zur selben Ordnung, so ergibt sich schließlich:

$$P_x = \frac{1}{\sqrt{2\pi \sum_1^s{}_i \sum_1^\omega{}_k p_{ik} S_{ik}^2}}\, e^{-\frac{x^2}{2\Sigma\Sigma p_{ik} S_{ik}^2}}.$$

Mit Hilfe dieser Elementarwahrscheinlichkeit findet sich die Wahrscheinlichkeit $\mathfrak{W}(-\varrho, \varrho)$, daß x in das Intervall $(-\varrho, \varrho)$ falle:

$$\mathfrak{W}(-\varrho, \varrho) = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt, \tag{2}$$

wobei zwischen γ und ϱ die Beziehung besteht:

$$\varrho = \gamma \sqrt{2 \sum_1^s{}_i \sum_1^\omega{}_k p_{ik} S_{ik}^2}. \tag{3}$$

Wählt man $\mathfrak{W}(-\varrho, \varrho)$ genügend groß — und in dieser Wahl liegt die unvermeidliche Willkür —, so ergibt das zugehörige ϱ jene Grenze, welche der eventuelle Gesamtverlust aus dem Massenspiel mit der angenommenen Wahrscheinlichkeit nicht überschreitet, und in diesem *absoluten Risiko*, wie es Küttner nennt, erblickt er das korrekte Ausmaß für den Verlustdeckungsfonds. Wie man sieht, ist

---

1) Diese Beschränkung stützt sich auf die Voraussetzung, die Zahlungseinheit sei so groß gewählt, daß sämtliche S_{ik} genügend kleine echte Brüche sind.

dieses absolute Risikos nicht eine feste, sondern eine schwankende Größe, die davon abhängt, bei welcher Wahl von $\mathfrak{W}(-\varrho, \varrho)$ man sich über die Zulänglichkeit des Fonds beruhigen will. Küttner selbst schlägt $\gamma = 2{,}33$ vor, wofür $\mathfrak{W}$ etwas wenig über $\frac{999}{1000}$ beträgt; das absolute Risiko wird dann[1])

$$\varrho = 3{,}3 \sqrt{\sum_1^s{}^i \sum_1^\omega{}^k p_{ik} S_{ik}^2}\,. \tag{4}$$

Nun aber ist, wie ein Blick auf die Gleichung (1) in Nr. 125 lehrt, der Wurzelausdruck nichts anderes als das dem vorliegenden Massenspiel entsprechende *mittlere Risiko M*, so daß man wieder dahin gelangt, den erforderlichen Verlustdeckungsfonds durch ein Vielfaches des mittleren Risiko (oder des durchschnittlichen, das ja zu dem mittleren in einem festen Verhältnis steht) zu bemessen in dem Bewußtsein, daß seine Zulänglichkeit nicht sicher, aber in hohem Grade wahrscheinlich ist, wie dies schon am Schlusse von Nr. 125 ausgeführt wurde.

Küttners Theorie führt also nur scheinbar zu einem neuen Risikobegriff; in Wirklichkeit ist sein absolutes Risiko nichts anderes als ein willkürlich festzusetzendes Vielfaches des mittleren Risikos.

§ 3. Die moralische Erwartung.

128. Die Hypothese von Daniel Bernoulli. Die Beurteilung vom Zufall abhängiger Unternehmungen auf Grundlage des Prinzips der mathematischen Erwartung richtet sich lediglich nach den möglichen Gewinnsummen und Verlusten und nach den Wahrscheinlichkeiten, mit welchen sie zu erwarten sind, und sieht von den persönlichen Verhältnissen der Beteiligten völlig ab; sie verdient deshalb die Bezeichnung einer objektiven Beurteilung.

In der Wirklichkeit jedoch hält jede Person die Gewinnaussichten und Verlusterwartungen ihren Verhältnissen gegenüber, und wie mannigfach diese auch sein mögen, die Vermögenslage der Person wird dabei immer eine Rolle spielen.

Daniel Bernoulli[2]) war der erste, welcher auf den Unterschied zwischen dem physischen Wert einer Geldsumme und dem mit ihr

1) Bei Küttner steht der Faktor 3,6, weil er auch eine die Verlustgefahr begünstigende Unsicherheit der Wahrscheinlichkeiten p_{ik} in Rechnung zieht.

2) Specimen Theoriae Novae de Mensura Sortis. Comm. Ac. Petropol. V, 1738. Deutsch mit Noten von A. Pringsheim und einer Einleitung von L. Fick. Leipzig, 1896.

verbundenen Vorteile, dem durch sie geschaffenen Nutzen, den er in Geld ausgedrückt als den relativen oder moralischen Wert dem physischen gegenüberstellte, und auf den Einfluß aufmerksam wurde, den das Vermögen der Person auf die Bildung dieses Wertes ausübt. Er stellte hierüber die folgende Hypothese auf:

Der aus einem beliebig kleinen Vermögenszuwachs resultierende Vorteil, oder sein moralischer Wert, ist dem Zuwachs selbst direkt, dem vorhandenen Vermögen umgekehrt proportional.

Erfährt also das momentane Vermögen x einer Person den Zuwachs dx, so ist der moralische Wert dieser Änderung

$$dy = k\frac{dx}{x}, \tag{1}$$

wobei k eine von den sonstigen Umständen der Person abhängige Konstante bedeutet. Hat sich das Vermögen von dem Anfangswerte a durch derlei kleine Änderungen in den Endbetrag x umgewandelt, so ist der moralische Wert dieser endlichen Änderung

$$y = k\int_a^x \frac{dx}{x} = kl\cdot\frac{x}{a}\cdot \tag{2}$$

Bernoulli hat über die Beurteilung des Vermögens einer Person solche Bestimmungen aufgestellt, welche die Fälle $a = 0$ und $a < 0$ ausschließen. So erblickt er Vermögen auch in der Fähigkeit eines nichts oder gar nur Schulden besitzenden Menschen, seine Existenz durch Arbeit oder selbst durch Betteln zu erhalten; denn ein solcher würde auch gegen eine gewisse Geldsumme, gegen Bezahlung der Schulden und ein darüber hinausgehendes Geschenk auf die Ausübung jener Tätigkeit nicht verzichten wollen.

Die Formel (2) gibt y positiv oder negativ, je nachdem ein Vermögenszuwachs $(x > a)$ oder eine Vermögensabnahme $(x < a)$ stattgefunden hat.

Zu der Hypothese selbst ist zu bemerken, daß ihr erster Teil, wonach der Vorteil dem Zuwachs selbst proportional ist, wohl nicht angefochten werden kann; auch daß der Vorteil mit der Größe des bereits vorhandenen Vermögens abnehme, dürfte in den meisten Fällen zutreffen; daß er aber dieser Größe geradezu umgekehrt proportional sei, ist ohne Zweifel eine willkürliche Festsetzung, für deren Zulässigkeit nur die Übereinstimmung ins Feld geführt werden kann, welche zwischen den aus der Hypothese abgeleiteten Resultaten und den Eingebungen des gemeinen Verstandes besteht.

Warum er annimmt, daß das Vermögen einer Person immer nur durch sukzessives Hinzutreten unendlich kleiner Elemente stetig sich vermehre, erklärt Bernoulli mit der Rücksichtnahme auf leichteres

Verständnis. Die Wirkung dieser Annahme liegt aber hauptsächlich in der Vereinfachung der Rechnungen.[1])

129. Die moralische Erwartung. Bezüglich der Beurteilung eines vom Zufall abhängigen Unternehmens auf Grund seiner Hypothese stellt Bernoulli die Regel auf, daß man aus den relativen oder moralischen Werten der einzelnen in Aussicht stehenden Gewinne und Verluste mittels der ihnen zukommenden Wahrscheinlichkeiten den Mittelwert zu bilden und sodann die diesem Mittelwert entsprechende Änderung des ursprünglichen Vermögens zu bestimmen habe; diese sei ein Wertmaß für das betreffende Unternehmen in Ansehung der betrachteten Person. Diese Größe ist es, welche später, von Laplace[2]), als *moralische Hoffnung* oder Erwartung bezeichnet worden ist.

Die Durchführung des entwickelten Gedankens stellt sich wie folgt dar.

Einer Person, deren Vermögen den Wert a hat, stehen die Vermögensänderungen $\alpha, \beta, \gamma, \cdots$ mit den bezüglichen Wahrscheinlichkeiten $p, q, r, \cdots$ bevor; Gewinne werden positiv, Verluste negativ in Rechnung gebracht; mit der Festsetzung, daß $p + q + r + \cdots = 1$ sein soll, wird ausgesagt, daß von den Änderungen nur eine wirklich eintreten kann und auch eintreten muß.

Die relativen Werte jener Änderungen sind

$$kl \cdot \frac{a+\alpha}{a}, \quad kl \cdot \frac{a+\beta}{a}, \quad kl \cdot \frac{a+\gamma}{a}, \quad \cdots;$$

der Mittelwert dieser Beträge ist

$$pkl \cdot \frac{a+\alpha}{a} + qkl \cdot \frac{a+\beta}{a} + rkl \cdot \frac{a+\gamma}{a} + \cdots$$

d. i.

$$kl \cdot \frac{(a+\alpha)^p (a+\beta)^q (a+\gamma)^r \cdots}{a};$$

ist h die Vermögensänderung, die eine diesem Mittelwerte gleiche relative Bedeutung besitzt, so besteht die Gleichung:

$$kl \cdot \frac{a+h}{a} = kl \cdot \frac{(a+\alpha)^p (a+\beta)^q (a+\gamma)^r \cdots}{a},$$

aus welcher sich die moralische Erwartung

$$h = (a+\alpha)^p (a+\beta)^q (a+\gamma)^r \cdots - a \tag{3}$$

ergibt.

---

1) Betrachtungen über die Bedeutung dieser Annahme und ihre Vergleichung mit einer andern, von Buffon herrührenden, wonach die moralische Bedeutung einer *endlichen* Vermögensänderung ihr selbst direkt, dem durch sie geänderten Vermögen umgekehrt proportional angenommen wird, stellt H. E. Timerding in dem Aufsatze „Die Bernoullische Werttheorie" an; Zeitschr. f. Mathem. u. Phys., 47 (1902), p. 325—328.

2) Théorie analyt. d. prob., p. 432.

Sind die Beträge α, β, γ, $\cdots$ sehr klein im Vergleich zu a, und beschränkt man sich bei der Entwicklung von

$$\frac{h}{a} = \left(1+\frac{\alpha}{a}\right)^p \left(1+\frac{\beta}{a}\right)^q \left(1+\frac{\gamma}{a}\right)^r \cdots - 1$$

auf die ersten Potenzen der sehr kleinen echten Brüche $\frac{\alpha}{a}, \frac{\beta}{a}, \frac{\gamma}{a}, \cdots$, so wird

$$\frac{h}{a} = p\frac{\alpha}{a} + q\frac{\beta}{a} + r\frac{\gamma}{a} + \cdots$$

und

$$h = p\alpha + q\beta + r\gamma + \cdots;$$

d. h. je größer das Vermögen der Person im Vergleich zu den erwarteten Gewinnsummen und Verlusten ist, um so näher kommt die moralische Hoffnung der mathematischen.

130. Folgerungen aus dem Begriff der moralischen Erwartung.

1. Auch ein nach den Regeln der mathematischen Hoffnung geregeltes Unternehmen ist für jeden der Beteiligten von moralischem Nachteil.

Betragen in einem Glücksspiele oder einer Wette die Einsätze der Partner α, β, sind p, $q = 1 - p$ für sie die Wahrscheinlichkeiten zu gewinnen, und ist a das Vermögen des ersten Partners, so ist seine moralische Hoffnung

$$h = (a+\beta)^p (a-\alpha)^q - a.$$

Ist nun das Spiel mathematisch geregelt, also $p\beta = q\alpha$, so kann p durch $\frac{\alpha}{\alpha+\beta}$, q durch $\frac{\beta}{\alpha+\beta}$ ersetzt und deshalb

$$h = [(a+\beta)^\alpha (a-\alpha)^\beta]^{\frac{1}{\alpha+\beta}} - a$$

geschrieben werden. Nun ist aber das geometrische Mittel einer Anzahl von Größen kleiner als das arithmetische[1]), folglich

$$[(a+\beta)^\alpha (a-\alpha)^\beta]^{\frac{1}{\alpha+\beta}} < \frac{\alpha(a+\beta) + \beta(a-\alpha)}{\alpha+\beta} = a,$$

daher

$$h < 0.$$

2. Ein Kaufmann, dessen Vermögen a ist, realisiert eine Summe α, wenn ein Warenschiff glücklich einläuft, was mit der Wahrscheinlich-

1) Man denke sich eine Zahl S in n Teile $x, y, z, \cdots$ geteilt und bilde deren Produkt $xyz\cdots$; sobald in diesem zwei Faktoren ungleich sind, etwa x, y, kann man es dadurch vergrößern, daß man diese ungleichen Faktoren jeden durch ihr arithmetisches Mittel $\frac{x+y}{2}$ ersetzt, weil $\left(\frac{x+y}{2}\right)^2 > xy$ ist; man kann so mit der Vergrößerung des Produktes fortfahren, solange zwei un-

keit p zu erwarten sein möge. Seine Verlusterwartung ist $(1-p)\alpha$, und gegen eine ihr gleichkommende Prämie möge er sich bei einer Anstalt gegen den Verlust versichern. Daß diese Versicherung für ihn von Vorteil ist, zeigt folgende Überlegung.

Schließt er sie ab, so ist der Vermögenszuwachs

$$\alpha-(1-p)\alpha=p\alpha$$

sicher, und dieser hat für ihn den moralischen Wert

$$kl\cdot\frac{a+p\alpha}{a};$$

schließt er die Versicherung nicht ab, so hat er den Vermögenszuwachs α mit der Wahrscheinlichkeit p zu erwarten, und der moralische Wert hiervon ist

$$plk\cdot\frac{a+\alpha}{a};$$

nun ist aber

$$\frac{a+p\alpha}{a}>\left(\frac{a+\alpha}{a}\right)^p,$$

weil[1])

$$a+p\alpha>(a+\alpha)^p a^{1-p}$$

ist; die Versicherung ist also tatsächlich von Vorteil.

Sie bleibt es auch dann noch, wenn zu der Versicherungsprämie $(1-p)\alpha$ ein Zuschlag z kommt, solange dieser eine gewisse Grenze nicht überschreitet. Diese Grenze bestimmt sich aus der Gleichung

$$\frac{a+p\alpha-z}{a}=\left(\frac{a+\alpha}{a}\right)^p,$$

welche erfüllt sein muß, damit die nunmehr geltenden relativen Werte

---

gleiche Faktoren darin vorkommen; den größten Wert nimmt aber das Produkt an, wenn alle Faktoren gleich und $=\frac{x+y+z+\cdots}{n}$ werden; somit ist

$$xyz\cdots<\left(\frac{x+y+z+\cdots}{n}\right)^n$$

und daraus

$$(xyz\cdots)^{\frac{1}{n}}<\frac{x+y+z+\cdots}{n}. \qquad (4)$$

Im obigen Falle besteht das Produkt aus α Faktoren $a+\beta$ und β Faktoren $a-\alpha$.

1) Macht man in (4) m Faktoren gleich $a+\alpha$, $n-m$ Faktoren gleich a und richtet die Zahlen m, n so ein, daß $\frac{m}{n}=p$ wird, so ergibt sich

$$[(a+\alpha)^m a^{n-m}]^{\frac{1}{n}}<\frac{m(a+\alpha)+(n-m)a}{n}$$

d. h.

$$(a+\alpha)^p a^{1-p}<a+p\alpha.$$

$$kl\frac{a+p\alpha-z}{a} \text{ und } pkl\frac{a+\alpha}{a}$$

einander gleich seien; sie gibt

$$z = a + p\alpha - (a+\alpha)^p a^{1-p}.$$

Hätte z. B. ein Kaufmann, dessen Vermögen $a = 5043$ Frcs. beträgt, eine Schiffsladung im Werte $\alpha = 10000$ Frcs. unterwegs, und betrüge die Wahrscheinlichkeit, mit welcher Schiffe der betreffenden Linie glücklich einlaufen, $\frac{19}{20}$, so wäre die mathematische Versicherungsprämie $\frac{1}{20} \cdot 10000 = 500$ Frcs.; ohne daß die Versicherung moralisch nachteilig wird, verträgt diese Prämie einen Zuschlag bis zur Höhe

$$z = 5043 + \frac{19}{20} \cdot 10000 - 15043^{\frac{19}{20}} \cdot 5043^{\frac{1}{20}},$$

d. i.

$$z = 300 \text{ Frcs. oder } 60\% \text{ der Prämie.}^{1)}$$

3. Es ist moralisch vorteilhafter, eine gefährdete Summe auf mehrere gleichartige Gefahren zu verteilen, statt sie ungeteilt *einer* solchen Gefahr auszusetzen.[2])

Wir begnügen uns, dies für die gleichmäßige Aufteilung auf zwei Gefahren nachzuweisen.

Erwartet ein Kaufmann, dessen Vermögen a ist, ein Gut von dem Werte α mit einem Schiff, dessen glückliches Eintreffen mit der Wahrscheinlichkeit p zu erhoffen ist, so ist der moralische Wert dieser Erwartung proportional

$$l \cdot (a+\alpha)^p a^q;$$

hätte er die Ladung auf zwei Schiffe dieser Art zu gleichen Teilen verteilt, so hätte er mit der Wahrscheinlichkeit p^2 das Eintreffen des ganzen Gutes, mit der Wahrscheinlichkeit $2pq$ das Eintreffen der Hälfte und mit der Wahrscheinlichkeit q^2 den Verlust des Ganzen zu erwarten; der moralische Wert dieser Erwartung ist proportional dem

$$l \cdot (a+\alpha)^{p^2}\left(a+\frac{\alpha}{2}\right)^{2pq} a^{q^2}.$$

Nun ist tatsächlich

$$(a+\alpha)^{p^2}\left(a+\frac{\alpha}{2}\right)^{2pq} a^{q^2} > (a+\alpha)^p a^q;$$

---

1) Vgl. § 15 des Specimen.

2) Crofton (Encycl. Brit. XIX, 1885, Art. 25) bemerkt, die Sentenz „es ist nicht gut, alle Eier in ein Nest zu legen" sei eine sprichwörtliche Anerkennung dieses Prinzips.

denn aus diesem Ansatze folgt:

$$(a+\alpha)^{p^2-p}\left(a+\frac{\alpha}{2}\right)^{2pq}a^{q^2-q}>1$$

oder

$$(a+\alpha)^{-pq}\left(a+\frac{\alpha}{2}\right)^{2pq}a^{-pq}>1$$

oder

$$\frac{\left(a+\frac{\alpha}{2}\right)^{2pq}}{(a+\alpha)^{pq}a^{pq}}>1,$$

und dies ist richtig, weil wegen

$$\left(a+\frac{\alpha}{2}\right)^2>(a+\alpha)a$$

auch

$$\frac{\left(a+\frac{\alpha}{2}\right)^2}{(a+\alpha)a}>1$$

ist.[1])

131. Das Petersburger Problem. Daniel Bernoulli wurde zur Aufstellung der Lehre von der moralischen Hoffnung durch eine spezielle Aufgabe geführt, die ihm sein Oheim Nicolaus Bernoulli vorlegte, nachdem er sie früher schon (1713) Montmort[2]) gestellt hatte. Es ist dies das *Petersburger Problem*, so genannt wegen seiner Veröffentlichung in den Akten der Petersburger Akademie, das wie folgt lautet: „Peter wirft eine Münze so lange, bis sie Wappen zeigt; geschieht dies beim ersten Wurfe, so hat er Paul 1 Dukaten zu geben; wenn erst beim zweiten Wurfe, 2 Dukaten; wenn erst beim dritten Wurfe, 4 Dukaten, und so bei jedem späteren Wurfe doppelt so viel, als wenn es im vorangehenden geschehen wäre. Wie groß ist die mathematische Erwartung Pauls?"

Die Rechnung gibt, da die Wahrscheinlichkeit der bezeichneten Fälle $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots$ ist, für die Erwartung Pauls

$$\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2+\frac{1}{8}\cdot 4+\cdots,$$

wenn, wie es dem Wortlaute der Aufgabe entspricht, das Spiel unbegrenzt fortsetzbar gedacht wird, einen unendlichen Wert, fordert daher auch einen unendlich großen Einsatz von Paul, und dies steht, wenn es überhaupt einen vernünftigen Sinn hat, mit der Erkenntnis im Widerspruche, daß niemand auf die Gewinnaussicht, welche das

1) Laplace, Théor. analyt. d. prob., p. 435ff., hat den Nachweis für eine beliebige Anzahl gleicher Teile gegeben.

2) Essai d'analyse sur les jeux de hazard (1714), p. 402.

Spiel eröffnet, einen nur einigermaßen erheblichen Einsatz leisten möchte.

D. Bernoulli suchte nun diesen Widerspruch durch die Anwendung des Begriffes der moralischen Hoffnung zu lösen, indem er zeigte, daß diese bei einem noch so großen Vermögen Pauls einen endlichen Wert besitzt und daher nur einen beschränkten Einsatz als vernünftig erscheinen läßt.

Die moralische Hoffnung Pauls, wenn sein Vermögen a ist, drückt sich nämlich aus durch

$$h = (a+1)^{\frac{1}{2}}(a+2)^{\frac{1}{4}}(a+4)^{\frac{1}{8}}\cdots - a,$$

und dieser Wert ist endlich, weil das unendliche Produkt

$$P(a) = \prod_1^\infty (a+2^{n-1})^{\frac{1}{2^n}}$$

für jeden Wert von a konvergiert. Zunächst gilt dies von

$$P(0) = \prod_1^\infty 2^{\frac{n-1}{2^n}},$$

weil

$$l\cdot P(0) = l\cdot 2\sum_1^\infty \frac{n-1}{2^n} = \frac{l\cdot 2}{2^2}\left(1+\frac{2}{2}+\frac{3}{2^2}+\frac{4}{2^3}+\cdots\right) = l\cdot 2,$$

wenn man bemerkt, daß die Reihe die Entwicklung von $\left(1-\frac{1}{2}\right)^{-2}$ vorstellt. Da weiter

$$\frac{P(a)}{P(0)} = \frac{1}{2}P(a) = \prod_1^\infty\left(1+\frac{a}{2^{n-1}}\right)^{\frac{1}{2^n}} < \prod_1^\infty (1+a)^{\frac{1}{2^n}}$$
$$= (1+a)^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots} = 1+a$$

ist, so hat man $P(a) < 2(1+a)$, wodurch die Konvergenz bewiesen ist.[1])

Bezeichnet man mit x die Grenze, bis zu welcher Paul mit seinem Einsatz gehen darf, ohne einen moralischen Nachteil zu haben, so bestimmt sich x aus der Gleichung:

$$(a-x+1)^{\frac{1}{2}}(a-x+2)^{\frac{1}{4}}(a-x+4)^{\frac{1}{8}}\cdots - a = 0;$$

setzt man $a - x = a'$, so wird

$$x = (a'+1)^{\frac{1}{2}}(a'+2)^{\frac{1}{4}}(a'+4)^{\frac{1}{8}}\cdots - a';$$

da aber bei einigermaßen großem a' das Verhältnis von x zu a' klein ausfällt, a' also wenig verschieden ist von a, so ist genau genug

1) Pringsheim l. c., Note 12.

$$x = (a+1)^{\frac{1}{2}}(a+2)^{\frac{1}{4}}(a+4)^{\frac{1}{8}}\cdots - a$$

oder für die Rechnung zweckmäßiger

$$\frac{x}{a} = \left(1+\frac{1}{a}\right)^{\frac{1}{2}}\left(1+\frac{2}{a}\right)^{\frac{1}{4}}\left(1+\frac{4}{a}\right)^{\frac{1}{8}}\cdots - 1.$$

Man findet beispielsweise, wenn $a = 100$, für x den Wert 4,36 (es genügen die ersten 15 Faktoren), wenn $a = 100$, $x = 6$.

Es entsteht aber die Frage, ob es erst der Theorie der moralischen Hoffnung bedurfte, um das Paradoxe an der Lösung zu beheben. Diese Frage ist zu verneinen. Der Grund des sinnlosen Resultates: ein unendlich großer Einsatz — liegt schon in der Stellung der Aufgabe, welche ein Spiel betrifft, das möglicherweise niemals endet und möglicherweise von Peter die Auszahlung einer Gewinnsumme fordert, die größer als jeder noch so große Betrag, die in der Sprache der Analysis unendlich ist; das aber sind bloße Worte ohne realen Sinn. Die Aufgabe enthält also von Hause aus eine Bestimmung, die in ihren Konsequenzen zu Ungereimtheiten führt.

Beschränkt man das Spiel auf eine bestimmte Anzahl von Würfen, z. B. n, so hat es zunächst nur dann einen vernünftigen Sinn, wenn Peters Vermögen auch die Auszahlung des höchsten möglichen Gewinnes 2^{n-1} gestattet. Der Einsatz bestimmt sich dann mit

$$\frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \cdots \frac{1}{2^n}\cdot 2^{n-1} = \frac{n}{2},$$

und das Paradoxe zum mindesten ist verschwunden. Will man beispielsweise annehmen, Peter besitze 1 Million Dukaten, so kann er noch $2^{19} = 524068$ auszahlen, sich also auf 20 Würfe einlassen, wofür ihm Paul einen Einsatz von 10 Dukaten zu leisten hätte.

Aber auch damit ist nicht alles Bedenkliche beseitigt, wenn man erwägt, daß Wappen erst im fünften Wurfe erscheinen dürfte, damit Paul durch die Annahme des Spiels keinen Verlust erleide; die Wahrscheinlichkeit hierfür ist aber nur $\frac{1}{32}$. Hier kommt nun der wesentliche Umstand zur Geltung, daß die mathematische Hoffnung für den Einzelfall keinen Anhalt bieten kann, daß ihre Bedeutung erst bei einem Massenspiel hervortritt, bei dem man erwarten darf, daß sich *alle* Möglichkeiten nahe im Verhältnis ihrer Wahrscheinlichkeiten zutragen, so daß der *durchschnittlich* auf ein Spiel entfallende Gewinn dem dafür geleisteten Einsatz nahe kommt. Gerade bei dem vorliegenden Spiel ist aber an eine solche Ausgleichung des Zufalls wegen der sehr großen Anzahl hierzu erforderlicher Spiele nicht zu denken.

Daß insbesondere der höchste Gewinn 2^{19} in s Spielen mindestens einmal eintrete, hat die Wahrscheinlichkeit[1])

$$1-\left(1-\frac{1}{2^{20}}\right)^s,$$

und soll diese auch nur den Wert $\frac{1}{2}$ erreichen, muß s mehr als 750000 betragen.[2])

132. Über die Bedeutung der Bernoullischen Hypothese. Die Lehre von der moralischen Hoffnung ist gewiß nicht geeignet, Unternehmungen, die auf zufälligen Ereignissen aufgebaut sind, zu regeln; dazu gibt das Prinzip der mathematischen Hoffnung die allein richtige Grundlage. Wenn es sich aber darum handelt, verschiedene Unternehmungen in ihren Wirkungen auf die Vermögenslage *einer* Person zu vergleichen, kann von ihr Gebrauch gemacht werden, und da führt sie, wie gezeigt worden, zu Resultaten, welche mit den Eingebungen des gemeinen Verstandes harmonieren; wäre ihre Leistung keine andere, als das mathematisch zu begründen, was die gesunde Überlegung an die Hand gibt, so hätte sie schon einen unbestreitbaren Wert.

Aber die Bedeutung der Theorie, insbesondere der ihr zugrunde liegenden Hypothese, reicht weiter. Diese Hypothese ist zur *Grundlage der modernen Wertlehre* geworden[3]).

Die ältere Wertlehre betrachtete den Wert eines Gutes als eine ihm innewohnende Eigenschaft. Diese Auffassung stammte daher, daß man sich bloß an den Tauschwert hielt und annahm, Güter, welche gegeneinander vertauscht werden, müßten gleichwertig sein, eine irrige Anschauung, da ja bei wirklicher Gleichwertigkeit kein Anlaß zum Tausche vorhanden wäre.

Die moderne Wertlehre, fast gleichzeitig begründet durch Jevons, Menger und Walras, betrachtet den Wert eines Gutes als eine Beziehung zwischen einem einzelnen wirtschaftenden Subjekt und dem Gute; nach dieser Auffassung läßt sich der Wertbegriff von der Vorstellung einer bestimmten Menge des Gutes nicht loslösen. Aus diesem Zusammenhange geht aber der für diese Theorie grundlegende Begriff des *Grenznutzens* hervor. Hierunter wird jener Nutzen verstanden, den der letzte zur Verfügung stehende Teil des Vorrates an einem Gute für ein bestimmtes wirtschaftendes Subjekt hat, und nach

---

1) S. Nr. 36.

2) Vgl. hiermit H. Bruns, Wahrscheinlichkeitsrechnung usw., p. 70—71; ferner A. Pringsheim, l. c., Note 10. Weiteres zur Geschichte des Petersburger Problems s. d. Verf. Abhandl. in Grunerts Arch. 77 (1881).

3) S. Ficks Einleitung zur deutschen Ausgabe des Specimen.

ihm bemißt sich der Wert einer Mengeneinheit des Gutes für das betreffende Subjekt.

Der Grundgedanke der neuen Auffassung, wonach der Wert einer neu hinzutretenden Einheit eines Gutes als eine abnehmende Funktion der Anzahl der bereits im Besitze der betreffenden Person befindlichen Einheiten sich darstellt, ist in der Bernoullischen Hypothese enthalten und hier zum ersten Male ausgesprochen worden; Bernoulli ging noch weiter, indem er über die Natur dieser Funktion eine bestimmte Annahme machte, die sich vielfach als zutreffend erwiesen hat.[1]) Hermann[2]) war wohl der erste Nationalökonom, der auf diese Bedeutung der Hypothese hingewiesen hat.

Sie ist indessen auch auf andere Gebiete übertragen worden, so von Fechner[3]) auf das Gebiet der Psychophysik, auf die Lehre von den Reizen und den durch sie hervorgerufenen Empfindungen; Friedr. Alb. Lange[4]) hat ihre Anwendbarkeit auf die Behandlung sozialer und politischer Probleme gezeigt.

---

1) Allgemeine Betrachtungen über diese sogenannte „Wertfunktion" und einige Anwendungen auf wirtschaftliche Fragen finden sich in dem unter Nr. 128 zitierten Aufsatze Timerdings.

2) Staatswirtschaftliche Untersuchungen. München 1832, p. 73.

3) Elemente der Psychophysik. Leipzig 1860, I, p. 230; II, p. 10.

4) Die Arbeiterfrage. Winterthur 1875, p. 113.

Zweiter Teil.

Ausgleichungsrechnung.

I. Abschnitt. Theorie der Beobachtungsfehler.

§ 1. Das Fehlergesetz.

133. Ziel der Ausgleichungsrechnung. Wenn zur Bestimmung einer oder mehrerer unbekannter Größen Messungen in der gerade notwendigen Anzahl ausgeführt werden, so bieten die Ergebnisse kein Mittel dar, um zu beurteilen, wie weit die aus ihnen abgeleiteten Werte der Unbekannten von den Beobachtungsfehlern beeinflußt sein können.

Sobald die Menge der ausgeführten Messungen das Maß des Notwendigen überschreitet, machen sich zwischen ihrem Ergebnisse *Widersprüche* bemerkbar, die ihren Grund in den Beobachtungsfehlern haben. Will man *alle* Beobachtungen zur Bestimmung der Unbekannten heranziehen, so stellt sich ein eigenartiges Problem ein: Die Ergebnisse der Messungen so zu kombinieren, daß die Bestimmung der Unbekannten eindeutig werde; dies erfordert, damit die Widersprüche zu bestehen aufhören, Abänderungen an den Beobachtungsresultaten oder, wie man dies zu bezeichnen pflegt, eine *Ausgleichung der Beobachtungen*.

Zur Erläuterung mögen die folgenden Beispiele dienen.

Sind zur Bestimmung einer physischen Größe mehrere unmittelbare oder *direkte* Messungen gemacht worden, so äußert sich der Widerspruch in der Ungleichheit der einzelnen Ergebnisse. Hat man diese Ergebnisse zur Gewinnung eines Wertes jener Größe irgendwie kombiniert, so erfordert die Aufhebung der Widersprüche, daß man an jedem Messungsresultate eine Korrektur anbringe, welche es dem aus der Kombination gewonnenen Werte der Unbekannten gleich macht.

Sind, um einen andern Fall zu betrachten, zwischen n von einem Punkte ausgehenden horizontalen Richtungen mehr als $n - 1$ Winkel gemessen worden, so werden sich einzelne der Winkel aus den gemessenen Winkeln auf verschiedene Weise ableiten lassen, und in der Nichtübereinstimmung der Resultate äußern sich die Widersprüche.

Wieder wird es sich darum handeln, alle Messungen zu einer eindeutigen Bestimmung der gegenseitigen Lage der n Strahlen zu kombinieren, und dies wieder wird mit einer Ausgleichung der Messungsergebnisse einhergehen.

Hat man in einem ebenen Dreieck alle drei Winkel gemessen, so werden die Ergebnisse der Messung vermöge der ihnen anhaftenden Fehler die von der Theorie geforderte Bedingung, zu 180° sich zu ergänzen, im allgemeinen nicht erfüllen; damit dieser Widerspruch mit der Theorie aufhöre, werden an den Messungsergebnissen Änderungen anzubringen sein, welche ihnen die Eigenschaft, die Winkel eines Dreiecks darstellen zu können, erst verleihen.

Wir kehren nun zu der allgemeinen Betrachtung zurück. Die in dem Problem verlangte Kombination der Beobachtungen hat zunächst nur die eine Forderung zu erfüllen, daß sie zu einer eindeutigen Bestimmung der Unbekannten führe. Dadurch ist aber die Methode ihrer Durchführung nach noch nicht gegeben, wird es auch dann nicht, wenn man die weitere Forderung hinzufügt, es sei unter allen Kombinationen die *vorteilhafteste* zu wählen. Denn dazu wäre ein mathematisch verwendbares Maß für die Beurteilung der Vorteilhaftigkeit einer Kombination erforderlich; ein solches aber läßt sich a priori nicht angeben.

Es liegt nahe, zu sagen, die vorteilhafteste unter allen Kombinationen sei diejenige, welche die Unbekannten mit den kleinstmöglichen Fehlern behaftet gibt. Da aber der Fehler einer Beobachtung selbst unbekannt ist, so wird man wenigstens Kenntnis darüber zu erlangen suchen, mit welcher Wahrscheinlichkeit er innerhalb bezeichneter Grenzen enthalten ist, um dann das Urteil über die größere oder geringere Vorteilhaftigkeit einer Methode auf die größere oder kleinere Wahrscheinlichkeit stützen zu können, daß der Fehler der Unbekannten, wenn sie nach dieser Methode abgeleitet wird, zwischen gewissen Grenzen eingeschlossen sei. Dadurch wäre das Problem auf das Gebiet der *Wahrscheinlichkeitsrechnung* übergeführt, und es bedürfte nur noch einer apriorischen Festsetzung über die Beziehung zwischen der Güte eines Messungsresultates und dem System der Fehler, welchen es unterworfen sein *kann*.

Dieser Gedankengang würde aber zu seiner strengen Durchführung erfordern, daß man die einzelnen störenden Ursachen auf ihre Wirksamkeit derart erforsche, um die Wahrscheinlichkeit angeben zu können, daß einer Beobachtung ein Fehler von bestimmter Größe zukomme. Eine solche Einsicht in die Entstehung der Fehler zu erlangen, erweist sich aber als unmöglich; daher ist auch eine exakte Begründung der Theorie der Kombination von Beobachtungen in dem gedachten Sinne, nämlich aus dem *wahren* Gesetz der Wahrscheinlichkeit der Beobachtungsfehler, untunlich.

Wohl aber ist es möglich, eine solche Theorie auf ein allgemeines *approximatives* Fehlergesetz zu stützen.

Neben der Bestimmung der Werte der unbekannten Größen geht noch eine andere Aufgabe der Ausgleichungsrechnung einher, die nämlich, aus den aufgetretenen Widersprüchen oder aus den an den Beobachtungsergebnissen zum Zwecke ihrer Ausgleichung anzubringenden Korrektionen auf die *Genauigkeit* der Beobachtungen selbst und der aus ihnen abgeleiteten Resultate zu schließen.[1])

134. Entstehung der Fehler und ihre Einteilung. Über die Entstehung der Beobachtungsfehler läßt sich zunächst mit Sicherheit nur aussagen, daß sie aus dem Zusammenwirken vieler Ursachen hervorgehen, die, wenigstens zum großen Teil, unabhängig voneinander sind. Je komplizierter der zur Messung verwendete Apparat, je mehr Einzelvorgänge zur Ausführung einer Beobachtung notwendig sind, um so größer die Anzahl der störenden Einflüsse.

Bei jeder Gattung von Beobachtungen wird sich eine Gruppe von Ursachen erkennen lassen, welche das Resultat in *systematischer* Weise beeinflussen, derart, daß ihre Wirkung in angebbarer Weise von gewissen Umständen abhängt. Solche Ursachen erzeugen *regelmäßige* oder *systematische*, unter gleichbleibenden Umständen insbesondere *konstante Fehler*. So wird bei der Messung einer geraden Linie durch sukzessives Auftragen eines Meßstabes ein systematischer Fehler daraus entspringen, daß der Meßstab nicht genau in die Richtung der Geraden gebracht wird, die Länge der Geraden wird dadurch immer zu groß gefunden; die wechselnde Temperatur des Meßstabes während der Messung zieht gleichfalls einen regelmäßigen Fehler nach sich, weil der Einfluß dieser Ursache sich mit der Temperatur systematisch ändert. An einem Winkelmeßinstrumente (Theodolith) wird jede bleibende Abweichung der einzelnen Bestandteile von den theoretischen Vorraussetzungen die Quelle systematischer Fehler sein. Es ist Sache des Beobachters, derlei systematisch wirkende Fehlerursachen, sei es durch Berichtigung des Instruments, sei es durch entsprechende Anordnung der Beobachtungen oder durch Berechnung des Einflusses wenigstens bis auf kleine, der Wahrnehmung sich entziehende Reste unschädlich zu machen. Systematische Fehler sind also der Hauptsache nach vermeidlich.

Dann bleibt immer noch eine meist große Anzahl von Ursachen, teils in den verwendeten Beobachtungsmitteln, teils in Mängeln des Wahrnehmungsvermögens, teils in äußeren die Messung begleitenden Umständen gelegen, deren Wirksamkeit insofern das Merkmal des

---

1) Eine gedrängte Zusammenstellung über die Begriffs- und Gedankenbildungen der Ausgleichungsrechnung, ihre Nomenklatur, Literatur u. a. hat S. Wellisch gegeben: Theoretische und historische Betrachtungen über die Ausgleichungsrechnung. Österr. Zeitschr. für Vermessungswesen, Wien 1907.

Zufälligen an sich trägt, als sich über die jeweilige Größe des Einflusses und darüber, ob er in dem einen oder andern Sinne von der Wahrheit ablenkt, eine bestimmte Aussage nicht machen läßt. Die aus solchen Ursachen entspringenden unvermeidlichen Fehler werden als *unregelmäßige* oder *zufällige Fehler* bezeichnet. Sie sind es, welche den Gegenstand der allgemeinen Fehlertheorie bilden.

Über das Auftreten der zufälligen Fehler lehrt die *Erfahrung* in großen Zügen zweierlei: daß positive und negative Fehler gleichen Betrages nahe gleich häufig und kleine Fehler häufiger auftreten als die größeren, so daß bei jeder Art von Beobachtungen Fehler über eine gewisse Größe hinaus — wenn man von groben *Irrungen* absieht — nicht zu erwarten sind.

Diese Tatsachen lassen sich aus der Annahme des Zusammenwirkens mehrerer unabhängiger Fehlerursachen unschwer erklären[1]).

Angenommen, es würden bei einer Art von Beobachtungen drei Ursachen tätig sein, welche mit gleicher Leichtigkeit folgende Fehler erzeugen:

1. Ursache:	-2,	-1,	0		
2. Ursache:		-1,	0,	1	
3. Ursache:		-1,	0,	1,	2, 3.

Durch ihre Kombination können die Fehler

$$-4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \qquad (\alpha)$$

im Resultate entstehen, und zwar beziehungsweise auf

$$1 \quad 3 \quad 6 \quad 8 \quad 9 \quad 8 \quad 6 \quad 3 \quad 1$$

verschiedene Arten; so kommt z. B. der Fehler 0 auf folgende Arten zustande:

$$\begin{array}{rrr} -2, & -1, & 3 \\ -2, & 0, & 2 \\ -2, & 1, & 1 \\ -1, & -1, & 2 \\ -1, & 0, & 1 \\ -1, & 1, & 0 \\ 0, & -1, & 1 \\ 0, & 0, & 0 \\ 0, & 1, & -1. \end{array}$$

Schon hier erkennt man die rasche Abnahme der Wahrscheinlichkeit und daher auch der zu erwartenden relativen Häufigkeit der Fehler mit der Zunahme des Betrages der letzteren.

---

1) P. Pizzetti, I fondamenti matematici per la critica dei risultati sperimentali. Atti della R. Università di Genova 1892, p. 121.

Wären bei einer Gattung von Beobachtungen acht unabhängige Ursachen wirksam, deren jede einen Fehler aus der Reihe (α) mit gleicher Leichtigkeit hervorbringen kann, so würden sich die $9^8 = 43,046.721$ möglichen Verbindungen der Einzelfehler auf die nun möglichen Werte des Resultatsfehlers von -32 bis $+32$ beiläufig wie folgt verteilen:

rund 2300000	auf den Fehler 0,					
1500000	auf jeden der Fehler	-5	und	$+5$,
950000	„ „ „ „	-10	„	$+10$,
300000	„ „ „ „	-15	„	$+15$,
50000	„ „ „ „	-20	„	$+20$,
3500	„ „ „ „	-25	„	$+25$,
36	„ „ „ „	-30	„	$+30$,
1	„ „ „ „	-32	„	$+32$.

Die große Anzahl der Entstehungsarten der kleinen Fehler gegenüber der geringen Menge bei den großen Fehlern erklärt sich daraus, daß jene nicht bloß aus kleinen Einzelfehlern, sondern auch aus den größeren und großen hervorgehen können, wenn diese sich vermöge ihrer Vorzeichen zum großen Teil tilgen, während die großen Resultatsfehler nur aus großen Einzelfehlern vorherrschend gleichen Zeichens entstehen können.

Aus dieser Betrachtung geht die Erkenntnis hervor, daß die *relative Häufigkeit*, mit welcher Fehler eines gewissen Größenintervalls in einer großen Anzahl von Beobachtungen zu erwarten sind, mit der Größe der Fehler in einem Zusammenhange steht.

Wir gehen nun daran, diesen Zusammenhang, das *Fehlergesetz*, approximativ darzustellen aus der Annahme, daß der Beobachtungsfehler sich aus einer großen Anzahl von *Elementarfehlern*, entspringend aus verschiedenen unabhängigen Ursachen, zusammensetze, deren jeder nur sehr kleiner absoluter Beträge fähig ist. Wir folgen dabei im wesentlichen dem Gedankengange M. W. Croftons[1]).

---

1) „On the proof of the law of errors of observations", Lond. Trans. 159 (1869) und „Probability", Encycl. Brit. 19 (1885). — Zur Geschichte der andern Ableitungen des Fehlergesetzes auf dieser Grundlage vgl. des Verf.s „Theorie der Beobachtungsfehler", 1891, p. 61—97, sowie dessen „Bericht" im Jahresber. d. deutschen Mathematiker-Vereinigung, VII (1899), p. 163ff. — Aus jüngster Zeit seien hier drei Arbeiten angeführt: F. Hausdorff (Leipziger Ber. 1901, p. 166—178) zeigt durch seine Ableitung die Stellung des nach Gauß benannten Fehlergesetzes, zu dem die folgenden Nummern als Schlußresultat hinführen, als einer Appoximation allgemeinerer Entwicklungen, die ein Resultat in geschlossener Form nicht ergeben; charakteristisch für diese Entwicklungen sind die von Hausdorff eingeführten *kanonischen Parameter*, aus Durchschnittswerten der Fehlerpotenzen gebildet. — A. Sommerfeld (Boltzmann-Festschrift 1904, p. 848—859) geht von Elementarfehlern aus, die innerhalb bestimmter Grenzen konstante Wahrscheinlichkeit haben. Bezeichnend für seine Arbeit ist die Anwen-

135. Ableitung des Fehlergesetzes aus der Hypothese der Elementarfehler. Es sei der Beobachtungsfehler x additiv zusammengesetzt aus den Elementarfehlern

$$\varepsilon', \ \varepsilon'', \ \varepsilon''' \cdots,$$

welche der Reihe nach

$$n', \ n'', \ n''' \cdots$$

verschiedene Werte annehmen können. Dann ist

$$x = \varepsilon' + \varepsilon'' + \varepsilon''' + \cdots$$

selbst $n'n''n''' \cdots$ verschiedener Werte fähig. Statt dieses Produktes wollen wir eine für die ganze Untersuchung festbleibende Zahl N nehmen, jedoch so, daß die relative Häufigkeit der verschiedenen Werte von x dadurch nicht verändert wird.

Diese relative Häufigkeit sei durch

$$z = f(x) \tag{1}$$

ausgedrückt in dem Sinne, daß $z\,dx$ die Menge der Fehler bedeutet, welche in das Intervall $(x, x + dx)$ fallen. Dabei werde $f(x)$ als eine *analytische* Funktion vorausgesetzt.

Diese Voraussetzung entspricht allerdings nicht der Wirklichkeit; denn diese schließt es in jedem besonderen Falle aus, daß die Fehlergröße x als eine *stetige* Variable angesehen werde, da wir vermöge unseres begrenzten Wahrnehmungsvermögens und der begrenzten Teilung an den Instrumenten das Messungsresultat nur in gewissen Abstufungen zu *konstatieren* imstande sind. Soll aber eine alle Fälle umfassende Theorie aufgestellt werden, so bleibt nichts übrig, als sich mit gewissen Annahmen zufrieden zu geben, welche die Durchführung der Rechnung ermöglichen, und zu diesen Annahmen gehört auch die, daß nicht nur x eine stetige Variable, sondern auch $f(x)$ eine stetige nach der Taylorschen Formel entwickelbare Funktion sei.

Für N ergibt sich der analytische Ausdruck

$$N = \int_g^G f(x)\,dx,$$

---

dung geometrischer Betrachtungen, die im zwei- und dreidimensionalen Gebiet beginnend zu höheren Räumen und schließlich zum „Raume von unendlich vielen Dimensionen" fortschreiten. — Eine sehr eingehende kritische Untersuchung über die verschiedenen Ableitungen des Fehlergesetzes aus Elementarfehlern hat F. Y. Edgeworth (Law of Error, Cambridge Philos. Trans. XX, 1904, p. 36—65, 113—130, Append. p. 131—141 und I—XIV) gegeben. Sie geht in wesentlichen Punkten über die früheren Arbeiten hinaus, so bezüglich des Grades der Approximation, bezüglich der Voraussetzung über das Gesetz, nach welchem sich die Elementarfehler zum Gesamtfehler vereinigen (die übliche Annahme ist die einfache Summation); es werden auch „mehrdimensionale" Fehler, d. h. Fehler in der Ebene und im Raume in Betracht gezogen.

wenn g die Summe der unteren, G die Summe der oberen Grenzen aller $\varepsilon^{(i)}$ vorstellt.

Es komme nun ein weiterer Elementarfehler ε hinzu, der fähig sein möge der n Werte $e', e'', \cdots$. Alle Fehler, welche aus Werten von x des Intervalls $(x, x+dx)$ durch Hinzufügung eines jeden Wertes von ε erzeugt werden, fallen in ein gewisses Intervall $(X, X+dX)$, wobei allgemein $X = x + \varepsilon$ ist; ihre Anzahl ist darstellbar durch

$$\{f(X-e') + f(X-e'') + \cdots\}\, dx,$$

ihre relative Häufigkeit, wenn man wieder auf die frühere Gesamtzahl N reduziert, durch

$$Z = \frac{1}{n}\{f(X-e') + f(X-e'') + \cdots\}.$$

Beschränkt man sich bei der Entwicklung in der Klammer mit Rücksicht darauf, daß die Beträge der Elementarfehler als sehr klein vorausgesetzt werden, auf die ersten und zweiten Potenzen von $e', e'', \cdots$, so stellt sich die relative Häufigkeit des Gesamtfehlers x nunmehr dar wie folgt:

$$Z = z - \frac{e' + e'' + e''' + \cdots}{n} z' + \frac{e'^2 + e''^2 + e'''^2 + \cdots}{2n} z'', \tag{2}$$

wobei z', z'' die erste und zweite Ableitung von $f(x)$ bedeuten.

Setzt man den mittleren Wert von ε:

$$\frac{e' + e'' + e''' + \cdots}{n} = \alpha, \tag{3}$$

den mittleren Wert seines Quadrates:

$$\frac{e'^2 + e''^2 + e'''^2 + \cdots}{n} = \beta, \tag{4}$$

so ist kürzer

$$Z = z - \alpha z' + \frac{\beta}{2} z''. \tag{2*}$$

Tritt ein weiterer Elementarfehler mit den charakteristischen Größen α_1, β_1 hinzu, so verwandelt er die relative Häufigkeit des Fehlers x in

$$Z_1 = Z - \alpha_1 Z' + \frac{\beta_1}{2} Z'';$$

da nun

$$\begin{aligned} Z &= z - \alpha z' + \frac{\beta}{2} z'', \\ Z' &= \phantom{z - {}} z' - \alpha z'' \cdots, \\ Z'' &= \phantom{z - z' - {}} z'' \cdots, \end{aligned}$$

so ist, wenn man bei Größen der zweiten Kleinheitsordnung stehen bleibt,

$$Z_1 = z - (\alpha + \alpha_1) z' + \frac{\beta + \beta_1 + 2\alpha\alpha_1}{2} z''$$

oder

$$Z_1 = z - (\alpha + \alpha_1) z' + \frac{\beta + \beta_1 - \alpha^2 - \alpha_1{}^2 + (\alpha + \alpha_1)^2}{2} z''. \tag{5}$$

Es gehen also *irgend zwei* Elementarfehler in das Gesetz des zusammengesetzten Fehlers nur in den Verbindungen $\alpha + \alpha_1$, $\beta + \beta_1$, $\alpha^2 + \alpha_1{}^2$ ein; folglich werden *alle* Elementarfehler in dem Häufigkeitsgesetz des resultierenden Fehlers nur in den Verbindungen

$$\begin{aligned} \alpha + \alpha_1 + \alpha_2 + \cdots &= a, \\ \beta + \beta_1 + \beta_2 + \cdots &= b, \\ \alpha^2 + \alpha_1{}^2 + \alpha_2{}^2 + \cdots &= c \end{aligned} \tag{6}$$

erscheinen, dieses Gesetz muß also die Form

$$z = F(x, a, b - c) \tag{7}$$

haben, da, wie in (5) zu bemerken, die Aggregate $\beta + \beta_1$ und $\alpha^2 + \alpha_1{}^2$ als Differenz auftreten.

Um zur Kenntnis der Funktion F zu gelangen, lasse man einen Elementarfehler hinzukommen, dessen Mittelwert da ist, während das mittlere Quadrat neben diesem außer acht gelassen werden kann. Dann ist einerseits die Anderung der Häufigkeit von x:

$$d_a z = \frac{\partial z}{\partial a} da,$$

andererseits laut (2*):

$$d_a z = -\frac{\partial z}{\partial x} da;$$

folglich

$$\frac{\partial z}{\partial x} = -\frac{\partial z}{\partial a}, \tag{8}$$

woraus hervorgeht, daß x und a in z nur in der Verbindung $x - a$ erscheinen, so daß z die Struktur haben wird:

$$z = F(x - a,\ b - c). \tag{7*}$$

Tritt ferner ein Elementarfehler hinzu, dessen mittlerer Wert Null ist, während sein mittleres Quadrat db beträgt, so drückt sich die Änderung der Häufigkeit von x einerseits durch

$$d_b z = \frac{\partial z}{\partial b} db,$$

andererseits vermöge (2*) durch

$$d_b z = \frac{1}{2} \frac{\partial^2 z}{\partial x^2} db$$

aus; daraus folgt, daß

$$\frac{\partial z}{\partial b} = \frac{1}{2}\frac{\partial^2 z}{\partial x^2}. \tag{9}$$

Man denke sich weiter alle Werte der Elementarfehler in dem Verhältnisse $1 : r$ verändert; dann verändern sich in gleichem Verhältnisse auch die Werte von x und a, während sich b und c im Verhältnis $1 : r^2$ verändern; drückt man die Tatsache aus, daß nun in dem Intervall $[rx, r(x + dx)]$ ebenso viele Fehlerwerte liegen, als früher in dem Intervall $(x, x + dx)$ gelegen waren, so entsteht die Gleichung:

$$F(x - a, b - c)dx = F[r(x - c), r^2(b - c)]rdx,$$

oder mit den Abkürzungen

$$x - a = \xi, \quad b - c = \eta$$

geschrieben:

$$F(r\xi, r^2\eta) = \frac{1}{r}F(\xi, \eta).$$

Mit $r = 1 + \delta$, unter δ eine sehr kleine Zahl verstanden, gibt die Entwicklung bis auf Größen der Ordnung δ:

$$z + \frac{\partial z}{\partial \xi}\delta\xi + 2\frac{\partial z}{\partial \eta}\delta\eta = z - \delta z,$$

woraus sich

$$\xi\frac{\partial z}{\partial \xi} + 2\eta\frac{\partial z}{\partial \eta} = -z \tag{10}$$

ergibt. Das allgemeine Integral dieser linearen partiellen Differentialgleichung lautet:

$$z = \eta^{-\frac{1}{2}}\psi(\xi^2\eta^{-1}). \tag{11}$$

Nun folgt aber aus $\eta = b - c$, daß

$$\frac{\partial z}{\partial \eta} = \frac{\partial z}{\partial b},$$

und weiter aus (9) und $\xi = x - a$, daß

$$\frac{\partial z}{\partial \eta} = \frac{1}{2}\frac{\partial^2 z}{\partial \xi^2};$$

setzt man dies in (10) ein, so wird

$$\eta\frac{\partial^2 z}{\partial \xi^2} + \xi\frac{\partial z}{\partial \xi} + z = 0;$$

die linke Seite dieser Gleichung ist aber der partielle Differentialquotient von $\eta\frac{\partial z}{\partial \xi} + \xi z$ in bezug auf ξ, und da er beständig $= 0$ ist, so muß

$$\eta \frac{\partial z}{\partial \xi} + \xi z = \chi(\eta)$$

lediglich Funktion von η sein. Führt man hierin die allgemeine Lösung (11) ein, so ergibt sich:

$$2\xi\eta^{-\frac{1}{2}}\psi'(\xi^2\eta^{-1}) + \xi\eta^{-\frac{1}{2}}\psi(\xi^2\eta^{-1}) = \chi(\eta).$$

Hiernach sollte eine Funktion von $\xi^2\eta^{-1}$ identisch sein einer Funktion von η; das kann nur stattfinden, wenn beide konstant sind, wenn also

$$\chi(\eta) = \eta \frac{\partial z}{\partial \xi} + \xi z = \text{const.};$$

da aber nach (11) ξ und $\frac{\partial z}{\partial \xi}$ zugleich verschwinden, muß auch const. $= 0$, also

$$\eta \frac{\partial z}{\partial \xi} + \xi z = 0 \tag{12}$$

sein.

Demnach hat man, mit der Abkürzung $\xi^2\eta^{-1} = u$, vermöge (11) und (12) zur Bestimmung von ψ die Gleichung:

$$2\psi'(u) + \psi(u) = 0,$$

aus der

$$\psi(u) = Ce^{-\frac{u}{2}}$$

folgt. Geht man nun von u auf ξ, η und von diesen auf die ursprünglichen Argumente x, a, b, c zurück, so ergibt sich als approximatives Gesetz der Häufigkeit des Fehlers x das folgende:

$$z = \frac{C}{\sqrt{b-c}}\, e^{-\frac{(x-a)^2}{2(b-c)}}. \tag{13}$$

Durch Division von $z\,dx$ mit der Anzahl N aller Fehlerwerte, welche sich durch das über diese Werte ausgedehnte Integral

$$\int z\,dx$$

darstellt, erhält man die Wahrscheinlichkeit eines in das Intervall $(x, x + dx)$ fallenden Fehlers; mit Rücksicht auf die außerordentlich rasche Abnahme von z mit wachsendem Betrage von x dürfen die Grenzen der Integration mit $-\infty$ und $+\infty$ angesetzt werden; dadurch ergibt sich:

$$N = \frac{C}{\sqrt{b-c}} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2(b-c)}}\, dx = C\sqrt{2} \int_{-\infty}^{\infty} e^{-t^2}\, dt = C\sqrt{2\pi}$$

und für die erwähnte Wahrscheinlichkeit der Ausdruck

$$\frac{1}{\sqrt{2\pi(b-c)}} e^{-\frac{(x-a)^2}{2(b-c)}} dx. \tag{14}$$

Den Koeffizienten von dx, d. i. die Funktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi(b-c)}} e^{-\frac{(x-a)^2}{2(b-c)}}, \tag{15}$$

bezeichnet man als das *Gesetz der Fehlerwahrscheinlichkeit* oder kurz als *Fehlergesetz*.

Bemerkt sei, daß die Differenz $b - c$ wesentlich positiv ist; denn nach (3) und (4) ist

$$\beta - \alpha^2 = \frac{e'^2 + e''^2 + e'''^2 + \cdots}{n} - \frac{(e' + e'' + e''' + \cdots)^2}{n^2}$$
$$= \frac{(1+1+1+\cdots)(e'^2 + e''^2 + e''' + \cdots) - (e' + e'' + e''' + \cdots)^2}{n^2}$$
$$= \frac{(e' - e'')^2 + (e' - e''')^2 + \cdots + (e'' - e''')^2 + \cdots}{n^2} > 0,$$

daher nach (6) auch

$$b - c > 0.$$

Setzt man zur Abkürzung

$$\frac{1}{2(b-c)} = h^2, \tag{16}$$

so wird in einfacherer Schreibweise

$$\varphi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2(x-a)^2}. \tag{15*}$$

Die Größe a bedeutet den mittleren Wert von x; denn, weil

$$x = \varepsilon' + \varepsilon'' + \varepsilon''' + \cdots$$

und $\alpha', \alpha'', \alpha''', \cdots$ die mittleren Werte von $\varepsilon', \varepsilon'', \varepsilon''', \cdots$ sind, so ist

$$a = \alpha' + \alpha'' + \alpha''' + \cdots$$

der mittlere Wert von x. Für $x = a$ wird $\varphi(x)$ ein Maximum; a ist somit der relativ wahrscheinlichste Beobachtungsfehler.

Die Wahrscheinlichkeit, daß der Beobachtungsfehler von diesem wahrscheinlichsten Betrage höchstens um δ nach auf- oder abwärts abweiche, drückt sich durch

$$P = \frac{h}{\sqrt{\pi}} \int_{a-\delta}^{a+\delta} e^{-h^2(x-a)^2} dx = \frac{2}{\sqrt{\pi}} \int_0^{h\delta} e^{-t^2} dt = \Phi(h\delta) \tag{17}$$

aus und kann bei bekanntem h mit Hilfe der Tafel I oder II bestimmt werden.

Über die Bedeutung des Parameters h wird später gesprochen werden.

136. Ableitung des Fehlergesetzes aus der Hypothese des arithmetischen Mittels. Es sei $\varphi(v)dv$ die apriorische Wahrscheinlichkeit, einen Fehler zwischen den Grenzen v und $v + dv$ zu begehen.

Wir betrachten nun den praktisch einfachsten und wegen seines häufigen Vorkommens wichtigen Fall, daß eine unbekannte physische Größe, X, n-mal unter — soweit die Wahrnehmung zu urteilen gestattet — gleichen Umständen (von demselben Beobachter, mit gleicher Sorgfalt, mit denselben Instrumenten u. dgl.) beobachtet worden sei, und daß man für sie die Werte $l_1, l_2, \cdots l_n$ erhalten habe.

Die Wahrscheinlichkeit a priori, daß n unabhängige Beobachtungen die Werte $l_1, l_2, \cdots l_n$ ergeben werden, ist proportional dem Produkte

$$\varphi(X - l_1)\,\varphi(X - l_2)\cdots\varphi(X - l_n); \tag{1}$$

dem analogen Produkte

$$\varphi(x - l_1)\,\varphi(x - l_2)\cdots\varphi(x - l_n) \tag{2}$$

proportional ist nach dem Satze von Bayes (Nr 100) auch die Wahrscheinlichkeit a posteriori, daß, wenn die Messung wirklich die Werte $l_1, l_2, \cdots l_n$ ergab, x der wahre Wert der gemessenen Größe sei.

Es entsteht nun die Frage: Welche Annahme über x ist auf dieser Erfahrungsgrundlage die vorteilhafteste?

Darüber kann aber erst entschieden werden, nachdem man ein theoretisches Prinzip aufgestellt hat. Eines der ältesten Prinzipe, von denen Gebrauch gemacht worden ist, besteht darin, jenen Wert der Unbekannten als den vorteilhaftesten zu erklären, der den Ausdruck (2) zum Maximum macht, mit andern Worten: *denjenigen Wert, welcher den bekannten Messungsergebnissen Fehler zuschreibt, für deren Existenz die größte aposteriorische Wahrscheinlichkeit besteht.* Diesen Gedanken hat Daniel Bernoulli[1]) zuerst ausgesprochen und Gauß[2]) in seiner ersten Publikation über den Gegenstand verwendet. Man kann den aus diesem Prinzip abgeleiteten Wert als den *wahrscheinlichsten Wert* der Unbekannten bezeichnen.

Hiernach wäre bei bekanntem φ dieser Wert aus der Gleichung

$$\frac{\varphi'(x - l_1)}{\varphi(x - l_1)} + \frac{\varphi'(x - l_2)}{\varphi(x - l_2)} + \cdots + \frac{\varphi'(x - l_n)}{\varphi(x - l_n)} = 0 \tag{3}$$

zu berechnen, die sich durch Nullsetzen des logarithmischen Differentialquotienten von (2) nach x ergibt.

Soll jedoch diese Gleichung zur Bestimmung der Funktion φ

---

1) Dijudicatio maxime probabilis plurium observationum discrepantium etc. Acta Ac. Petrop. 1778.

2) Theoria motus etc. 1809.

verwendet werden, so ist eine Annahme über x erforderlich. Als solche hat Gauß (l. c.) die zugrunde gelegt, *daß unter den hier vorausgesetzten Umständen das arithmetische Mittel der Beobachtungsresultate der vorteilhafteste, also der wahrscheinlichste Wert sei.*

Dann muß also die Gleichung

$$x = \frac{l_1 + l_2 + \cdots + l_n}{n} \tag{4}$$

oder die durch Umformung aus ihr abgeleitete

$$(x - l_1) + (x - l_2) + \cdots + (x - l_n) = 0 \tag{5}$$

eine Folge von (3) sein.

Setzt man

$$x - l_i = \lambda_i, \quad \frac{\varphi'(\lambda_i)}{\varphi(\lambda_i)} = \psi(\lambda_i),$$

so kommt dies darauf zurück, die Funktion ψ so zu bestimmen, daß mit

$$\psi(\lambda_1) + \psi(\lambda_2) + \cdots + \psi(\lambda_n) = 0$$

zugleich

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$$

bestehe.

Bildet man aus diesen Gleichungen, nachdem man sie differenziert hat, mittels des unbestimmten Multiplikators $2k$ die *eine* Gleichung:

$$[\psi'(\lambda_1) - 2k]d\lambda_1 + [\psi'(\lambda_2) - 2k]d\lambda_2 + \cdots + [\psi'(\lambda_n) - 2k]d\lambda_n = 0,$$

in welcher die Differentiale voneinander unabhängig sind, so ergibt sich, daß jede der eingeklammerten Differenzen verschwindet, daß also ψ der Differentialgleichung

$$\psi'(\lambda) = 2k$$

genügen müsse. Daraus folgt

$$\psi(\lambda) = 2k\lambda,$$

also

$$\frac{\varphi'(\lambda)}{\varphi(\lambda)} = 2k\lambda,$$

woraus durch neuerliche Integration

$$l \cdot \varphi(\lambda) = l \cdot C + k\lambda^2$$

und

$$\varphi(\lambda) = Ce^{k\lambda^2}$$

erhalten wird.

Da nach allgemeiner Erfahrung φ eine abnehmende Funktion ist, ist k notwendig negativ; ersetzt man es mit Rücksicht darauf durch $-h^2$, so wird

$$\varphi(\lambda) = Ce^{-h^2\lambda^2}.$$

Das über alle möglichen Werte des Fehlers ausgedehnte Integral von $\varphi(\lambda)d\lambda$ hat als Wahrscheinlichkeit dafür, daß der Fehler einen dieser Werte annehme, den Wert 1; folglich ist

$$C\int_{-\infty}^{\infty} e^{-h^2\lambda^2}d\lambda = \frac{C}{h}\int_{-\infty}^{\infty} e^{-t^2}dt = \frac{C\sqrt{\pi}}{h} = 1,$$

woraus $C = \frac{h}{\sqrt{\pi}}$.

Hiermit ist endgültig das Fehlergesetz

$$\varphi(\lambda) = \frac{h}{\sqrt{\pi}} e^{-h^2\lambda^2} \tag{6}$$

gefunden.

Gegenüber der in der vorigen Nummer unter (15*) gefundenen Form zeigt diese zwei Unterschiede: dort war x der *wahre* Fehler, also die Abweichung vom wahren Werte — hier ist λ die Abweichung von dem vorteilhaftesten Werte, also vom arithmetischen Mittel, der sogenannte *plausibelste* oder *scheinbare* Fehler; dort war x um a, d. i. um den mittleren Wert von x, vermindert; hier entfällt dies, weil vermöge (5) die Summe der λ, also auch der mittlere Wert von λ, Null ist.

Nimmt man (6) der Form nach auch als das Gesetz des wahren Fehlers an, der mit ε bezeichnet werden möge, so stellt

$$\varphi(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2} \tag{6*}$$

ein symmetrisches, (15*) dagegen, d. i.

$$\varphi_1(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2(\varepsilon - a)^2}, \tag{15*}$$

wenn $a \neq 0$, ein asymmetrisches Fehlergesetz dar. Das letztere geht also in das erste über, wenn der Durchschnittswert von ε, d. i.

$$a = \int_{-\infty}^{\infty} \varepsilon\,\varphi_1(\varepsilon)d\varepsilon, \tag{7}$$

verschwindet. Ein solches Verhalten hat Gauß als charakteristisches Merkmal rein zufälliger Fehler aufgefaßt, das Vorhandensein eines von Null verschiedenen a hingegen als Merkmal für die Existenz *einseitig* wirkender Fehlerursachen; in diesem Falle nennt er a den *konstanten Teil* der Beobachtungsfehler.

Das Gesetz (6*) stellt also gewissermaßen einen idealen Fall vor, darin bestehend, daß keine einseitig wirkenden Ursachen vorhanden sind oder daß sie sich, wenn vorhanden, zu einer völlig symmetrischen Verteilung der Fehler kompensieren. Die Wirklichkeit wird wohl in der Regel eine wenn auch nur schwache Asymmetrie aufweisen.

Die vorstehende Ableitung hat die Beweiskraft, welche ihr aus der Hypothese des arithmetischen Mittels entspringt. Daß das arithmetische Mittel der vorteilhafteste, insbesondere daß es der wahrscheinlichste Wert sei, ist unbeweisbar; aber es lassen sich mannigfache innere und äußere Gründe anführen, welche für die Wahl dieses Wertes sprechen.[1])

137. Das Präzisionsmaß. Setzt man in dem Fehlergesetz

$$\varphi(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2(\varepsilon - a)^2}$$

$a = 0$, so bedeutet das in (17) der vorigen Nummer gerechnete

$$P = \frac{2}{\sqrt{\pi}} \int_0^{h\delta} e^{-t^2} dt = \Phi(h\delta)$$

die Wahrscheinlichkeit, daß einer Beobachtung ein zwischen $-\delta$ und $+\delta$ liegender Fehler zukomme. Das Intervall 2δ verengt sich bei festem P in demselben Verhältnisse, als h wächst; je enger aber das Intervall, innerhalb dessen der Beobachtungsfehler mit gegebener Wahrscheinlichkeit zu erwarten ist, für desto genauer wird man die Art der Beobachtungen erklären.

Der Parameter h hängt also mit der Genauigkeit der Beobachtungen zusammen, deren Fehlerfrequenz durch das Gesetz dargestellt ist. Gauß hat die *Genauigkeit* geradezu der Größe h proportional gesetzt und diese darum als *Präzisionsmaß* bezeichnet.

Die erste Ableitung des Fehlergesetzes hat gezeigt, wie das Präzisionsmaß mit den Elementarfehlern zusammenhängt; nach Gleichung (16) ist

$$h = \frac{1}{\sqrt{2(b - c)}},$$

und darin ist b die Summe der mittleren Quadrate, c die Summe der Quadrate der mittleren Werte der einzelnen Elementarfehler.

138. Gesetz, welchem eine lineare Funktion unabhängiger Beobachtungsfehler folgt. Es sei

$$F = k + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n \tag{1}$$

eine lineare Funktion der direkt beobachteten Größen $X_1, X_2, \ldots X_n$; die Beobachtungsresultate $l_1, l_2, \ldots l_n$ mögen aus Messungsvorgängen stammen, deren Fehler dem allgemeinen Gesetze

$$\frac{h}{\sqrt{\pi}} e^{-h^2(x-a)^2}$$

1) Vgl. des Verfassers „Theorie der Beobachtungsfehler", 1891, p. 16—47.

folgen und deren Präzisionsmaße $h_1, h_2, \ldots h_n$ sind. Rechnet man statt F den Ausdruck

$$f = k + \alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n,$$

so entsteht die Frage, welchem Gesetz der dadurch begangene Fehler

$$F - f = \alpha_1 (X_1 - l_1) + \alpha_2 (X_2 - l_2) + \cdots + \alpha_n (X_n - l_n) \quad (2)$$

oder

$$\mathsf{E} = \alpha_1 \varepsilon_1 + \alpha_2 \varepsilon_2 + \cdots + \alpha_n \varepsilon_n \quad (3)$$

folgt. Dies kommt darauf hinaus, die Wahrscheinlichkeit zu bestimmen, daß E zwischen den Grenzen z und $z + dz$ liege; diese aber ist gegeben durch das Integral

$$\int\limits_{(n)} \frac{h_1 h_2 \ldots h_n}{(\sqrt{\pi})^n} e^{-[h_1^2(\varepsilon_1 - a_1)^2 + h_2^2(\varepsilon_2 - a_2)^2 + \cdots + h_n^2(\varepsilon_n - a_n)^2]} d\varepsilon_1 d\varepsilon_2 \ldots d\varepsilon_n,$$

ausgedehnt über dasjenige Gebiet, das durch die Relation

$$z \leqq \mathsf{E} \leqq z + dz$$

begrenzt wird.

Um während der Integration von dieser Grenzbeziehung unabhängig zu bleiben, wendet man den Kunstgriff an, den Ausdruck unter dem Integral mit einem Faktor zu multiplizieren, der für $\mathsf{E} = z$ den Wert 1 hat und in allen andern Fällen verschwindet; zu einem solchen *Diskontinuitätsfaktor* führt das Integral

$$\int\limits_{-\pi}^{\pi} e^{m u \sqrt{-1}} du,$$

das für jedes ganzzahlige m verschwindet und für $m = 0$ den Wert 2π hat. Setzt man

$$\frac{u}{\delta} = \theta,$$

unter δ eine sehr kleine Größe, unter θ eine neue Variable verstanden, so wird hiernach

$$\frac{\delta}{2\pi} \int\limits_{-\frac{\pi}{\delta}}^{\frac{\pi}{\delta}} e^{m\delta \cdot \theta \sqrt{-1}} d\theta$$

1 oder 0, je nachdem $m\delta$ Null oder nicht Null ist. Für $\delta = dz$ und $m\delta = \mathsf{E} - z$ ergibt sich der für den vorliegenden Fall taugliche Diskontinuitätsfaktor

$$\frac{dz}{2\pi} \int\limits_{-\infty}^{\infty} e^{(\mathsf{E} - z)\theta \sqrt{-1}} d\theta;$$

nach seiner Anbringung darf in bezug auf alle ε zwischen den Grenzen $-\infty$ und $+\infty$ integriert werden. Der Ausdruck für die verlangte Wahrscheinlichkeit wird dann:

$$\frac{dz}{2\pi}\frac{h_1 h_2 \ldots h_n}{(\sqrt{\pi})^n}\int_{-\infty}^{\infty} d\theta \int_{-\infty}^{\infty} d\varepsilon_1 \int_{-\infty}^{\infty} d\varepsilon_2 \ldots \int_{-\infty}^{\infty} e^{-\sum h_i^2(\varepsilon_i - a_i)^2 + \left(\sum \alpha_i \varepsilon_i - z\right)\theta\sqrt{-1}} d\varepsilon_n. \quad (4)$$

Die auf ε_i bezügliche Integration

$$\int_{-\infty}^{\infty} e^{-h_i^2(\varepsilon_i - a_i)^2 + \alpha_i \varepsilon_i \theta \sqrt{-1}} d\varepsilon_i,$$

wenn man den Exponenten in

$$-\left[h_i(\varepsilon_i - a_i) - \frac{\alpha_i \theta}{2h_i}\sqrt{-1}\right]^2 + a_i \alpha_i \theta \sqrt{-1} - \frac{\alpha_i^2 \theta^2}{4h_i^2}$$

umformt, gibt

$$\frac{\sqrt{\pi}}{h_i} e^{a_i \alpha_i \theta \sqrt{-1} - \frac{\alpha_i^2 \theta^2}{4h_i^2}};$$

folglich wird aus (4) nach Ausführung der Integrationen in bezug auf alle ε:

$$\frac{dz}{2\pi}\int_{-\infty}^{\infty} e^{\theta\sqrt{-1}\sum a_i \alpha_i - \frac{\theta^2}{4}\sum \frac{\alpha_i^2}{h_i^2} - z\theta\sqrt{-1}} d\theta.$$

Setzt man zur Abkürzung

$$\sum a_i \alpha_i = A, \qquad \sum \frac{\alpha_i^2}{h_i^2} = \frac{1}{H^2}, \quad (5)$$

und formt hierauf den Exponenten wie folgt um:

$$-\left[\frac{\theta}{2H} + H(z - A)\sqrt{-1}\right]^2 - H^2(z - A)^2,$$

so ergibt sich schließlich

$$\frac{H}{\sqrt{\pi}} e^{-H^2(z - A)^2} dz \quad (6)$$

als Ausdruck für die eingangs bezeichnete Wahrscheinlichkeit.

Folgen also die Fehler der einzelnen Beobachtungen dem Exponentialgesetz $\frac{h}{\sqrt{\pi}} e^{-h^2(x-a)^2}$, *so befolgt der Fehler von f ein Exponentialgesetz von derselben Form.*

Sind die Beobachtungen gleich genau, so daß $h_1 = h_2 = \cdots = h_n = h$ ist, so ist nach (5) die Präzision in der Bestimmung von f:

$$H = \frac{h}{\sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}}. \tag{7}$$

Ein besonderer Fall, der praktische Verwendung findet, soll hier angemerkt werden. Setzt man in (1)

$$k = 0, \quad \alpha_1 = 1, \quad \alpha_2 = -1, \quad \alpha_3 = \alpha_4 = \cdots = \alpha_n = 0,$$
$$X_1 = X_2 = X,$$

so gehen F in Null, f und E in die Differenz zweier gleich genauen unabhängigen Beobachtungen einer und derselben Größe über; Formel (7) gibt für H den Wert $\frac{h}{\sqrt{2}}$.

Die Differenzen gleich genauer Beobachtungen von der Präzision h verhalten sich also in ihrer Frequenz so, wie Beobachtungsfehler von der Präzision $\frac{h}{\sqrt{2}}$.

Das durch (6) ausgedrückte Gesetz hat auch für eine *beliebige* Funktion $F(X_1, X_2, \ldots X_n)$ bedingte Geltung, unter der Voraussetzung nämlich, daß die $l_1, l_2, \ldots$ anhaftenden Fehler $\varepsilon_1, \varepsilon_2, \ldots$ so klein seien, daß man Potenzen und Produkte derselben gegenüber den ersten Potenzen vernachlässigen, also mit genügender Schärfe

$$F = F(l_1 + \varepsilon_1, l_2 + \varepsilon_2, \cdots l_n + \varepsilon_n) = k + \alpha_1 \varepsilon_1 + \alpha_2 \varepsilon_2 + \cdots + \alpha_n \varepsilon_n \tag{8}$$

setzen darf, wobei

$$\begin{aligned} k &= F(l_1, l_2, \ldots l_n), \\ \alpha_i &= F'_{X_i}(l_1, l_2, \ldots l_n). \end{aligned} \tag{9}$$

Demnach bilden die Formeln (5) bis (9) die Grundlage für die Beurteilung der Genauigkeit in der Bestimmung einer Funktion direkt beobachteter unabhängiger Größen.

139. Der wahre Wert einer physischen Größe.[1]) Man spricht von dem *wahren Werte* einer physischen Größe in der stillschweigenden Voraussetzung, daß ein solcher in einer jede Unbestimmtheit ausschließenden Weise definiert sei. Dies wird in manchen Fällen *innerhalb der Grenzen unseres Wahrnehmungsvermögens* auch zutreffen; in *absolutem* Sinne, d. h. bei unbegrenzter Schärfe dieses Vermögens, wird eine physische Größe niemals vollkommen definiert sein.

Die Basis einer großen Triangulierung mag noch so sorgfältig festgelegt sein, geometrisch streng definiert ist sie nicht, weil ihre Endpunkte nicht Punkte im eigentlichen Sinne, sondern sichtbare Marken sind.

Der Winkel der Richtungen, welche von einem „Punkte" der Erdoberfläche nach zwei andern „Punkten" gehen, ist nicht unzwei-

1) Vgl. P. Pizzetti, l. c., p. 146 ff.

deutig definiert, weil sowohl der Ausgangspunkt wie auch der Zielpunkt Objekte von endlicher Ausdehnung sind und sein müssen, um wahrgenommen zu werden; auch das Einstellen auf ihre „Mitte“ läßt noch eine Unbestimmtheit übrig.

Die zwischen den Endflächen eines Meßstabes enthaltene Länge entbehrt bei genauem Zusehen auch der strengen Definition; mögen die Endflächen noch so fein bearbeitet sein, vollkommene und parallele Ebenen im geometrischen Sinne stellen sie schon vermöge der Diskontinuität der Materie nicht dar.

Man spricht von der Zenitdistanz eines Sternes an einem bestimmten Orte und stellt sich dabei vor, daß die Lotrichtung an diesem vollkommen definiert sei; bedenkt man aber, daß die Lotrichtung im Zusammenhange steht mit der Verteilung der Massen, so wird man bei der beständigen Veränderlichkeit dieser Verteilung streng genommen nur von der Lotrichtung in einem bestimmten Augenblicke sprechen können.

Bei der begrenzten Schärfe unserer Beobachtungsmittel, der natürlichen wie der künstlichen, entziehen sich aber Unsicherheiten in der Definition der zu messenden Größe und zeitliche Änderungen derselben, sobald sie unter ein gewisses Maß herabsinken, der Wahrnehmung vollständig, und es bleibt die Möglichkeit, für den wahren Wert eine auf die Erfahrung gegründete Definition aufzustellen. Es soll darunter *die Grenze* verstanden werden, *welcher sich der aus einer großen Zahl von Beobachtungen abgeleitete Mittelwert*[1]) *nähert, wenn die Zahl der Beobachtungen beständig wächst.*

Diese Definition setzt voraus, daß der mathematische Ausdruck des Mittelwertes aus n Beobachtungen mit wachsendem n tatsächlich gegen eine bestimmte Grenze konvergiere, der dadurch, daß sie lediglich von der physischen Größe und nicht auch von dem Beobachter, den Instrumenten abhängt, eine *absolute* Bedeutung zukommt.

Das Vorhandensein eines Grenzwertes allein wäre nicht ausreichend zur Definition des wahren Wertes. Denn unterlägen die Beobachtungen einer konstant wirkenden Fehlerquelle, so ergäbe sich nach deren Ausscheidung (z. B. nach Berichtigung des Instrumentes oder Wahl eines andern) sofort ein von dem früheren verschiedener Grenzwert, gegen die zuletzt ausgesprochene Forderung.

Für die Methode der Kombination der Beobachtungen ergibt sich aus diesen Betrachtungen das folgende Postulat: *„Der Ausdruck, welcher den vorteilhaftesten Wert einer physischen Größe aus einer Anzahl ausgeführter Beobachtungen bestimmt, muß so beschaffen sein, daß er mit beständigem Wachsen dieser Anzahl dem wahren Werte der Größe als Grenze sich nähert.“*

---

1) Unter „Mittelwert“ wird hier der nach einer bestimmten Methode gewonnene vorteilhafteste Wert der unbekannten Größe verstanden.

Dies erfordert, daß in dem allgemeinen Fehlergesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2(\varepsilon - a)^2}$$

die Größe a, von Gauß[1]) der *konstante Teil des Fehlers* genannt, entweder im voraus, vor Anstellung der Beobachtungen, bekannt oder gleich Null sei.

Angenommen nämlich, zur Bestimmung einer unbekannten Größe X seien n Beobachtungen, deren Fehler dem obigen Gesetze folgen, ausgeführt worden und hätten die Resultate $l_1, l_2, \ldots l_n$ ergeben. Durch $f(l_1, l_2, \ldots l_n)$ sei die Rechnungsvorschrift dargestellt, welche den vorteilhaftesten Wert der Unbekannten liefern soll, so daß

$$x = f(l_1, l_2, \ldots l_n)$$

ist. Die erste Forderung, welche die Funktion f zu erfüllen hat, ist die, daß ihr Wert X sei, wenn man für die fehlerhaften Beobachtungsergebnisse X selbst einsetzt, daß also

$$X = f(l_1 + \varepsilon_1, \; l_2 + \varepsilon_2, \ldots l_n + \varepsilon_n).$$

Entwickelt man die rechte Seite unter der Voraussetzung sehr kleiner ε (vgl. Schluß der vorigen Nummer), so ergibt sich für den Fehler in der Bestimmung x die Darstellung:

$$X - x = \alpha_1 \varepsilon_1 + \alpha_2 \varepsilon_2 + \cdots + \alpha_n \varepsilon_n,$$

und dieser Fehler befolgt das Gesetz

$$\frac{H}{\sqrt{\pi}} e^{-H^2(z - A)^2},$$

wobei

$$H = \frac{h}{\sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}}, \qquad A = a(\alpha_1 + \alpha_2 + \cdots + \alpha_n).$$

Soll mit wachsendem n sich x der Größe X nähern, so muß notwendig der Fehler 0 die größte Häufigkeit haben; dies ist aber nur dann der Fall, wenn $A = 0$, was wiederum die Gleichung $a = 0$ nach sich zieht, weil das Nullwerden von $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ nicht vorausgesetzt werden darf.[2]) Des weiteren ist erforderlich, daß H

---

1) Theoria combinationis I (1821), Art. 5. Deutsch von A. Börsch und P. Simon, Berlin 1887. Vgl. auch Nr. 137.

2) Es ist nämlich für ein sehr kleines δ bis auf die Ordnung dieser Größe $f(l_1 + \delta, \; l_2 + \delta, \ldots l_n + \delta) = x + (\alpha_1 + \alpha_2 + \cdots + \alpha_n)\delta$; wäre $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 0$, so hätte man

$$f(l_1, l_2, \ldots l_n) = x = f(l_1 + \delta, l_2 + \delta, \ldots l_n + \delta),$$

d. h. an dem vorteilhaftesten Werte würde durch Abänderung sämtlicher Beobachtungen um δ keine Änderung hervorgebracht.

mit n beständig wachse, daß also die Summe $\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2$ beständig abnehme und sich der Null als Grenze nähere.

Der Fall, daß a nicht Null, aber bekannt sei, kann auf den früheren zurückgeführt werden. Schreibt man X in der Form:

$$X = f(l_1 - a + \varepsilon_1 + a, \; l_2 - a + \varepsilon_2 + a, \ldots l_n - a + \varepsilon_n + a),$$

entwickelt nach den als sehr klein vorausgesetzten Größen $\varepsilon_1 + a, \varepsilon_2 + a, \cdots$, und setzt

$$f(l_1 - a, \; l_2 - a, \; \cdots l_n - a) = x',$$

so wird

$$X - x' = \alpha_1' \varepsilon_1 + \alpha_2' \varepsilon_2 + \cdots + \alpha_n' \varepsilon_n + A',$$

wo A' wie vorhin die Bedeutung $a(\alpha_1' + \alpha_2' + \cdots + \alpha_n')$ hat; der Fehler in x' befolgt nun das Gesetz

$$\frac{H'}{\sqrt{\pi}} e^{-H'^2(\varepsilon + A' - A')^2}$$

d. i.

$$\frac{H'}{\sqrt{\pi}} e^{-H'^2 \varepsilon^2},$$

was mit der früheren Voraussetzung $A = 0$ übereinstimmt. Die Rechnungsvorschrift, welche für $a = 0$ zu dem wahren Werte als Grenze führt, leistet dies also auch bei $a \neq 0$, wenn man das als bekannt vorausgesetzte a von jeder Beobachtung subtrahiert.

§ 2. Genauigkeitsmaße.

140. Das Fehlerrisiko. Man kann die Bestimmung einer Größe durch Beobachtungen einem Spiele vergleichen, bei dem nur Verluste möglich sind; denn jeder Fehler, ob positiv oder negativ, bedeutet als Abweichung von der Wahrheit einen Verlust. Der mit einem Spiele verbundenen Verlustgefahr entspricht hier die Befürchtung eines Fehlers in dem Resultate, von deren Ausmaß die *Genauigkeit* abhängen wird, die man der gewonnenen Bestimmung zuschreibt.

Wie nun bei dem Spiele in dem Risiko ein Maß für die Gefährlichkeit desselben erkannt worden ist, kann eine in analoger Weise gebildete Größe bei Beobachtungen als Maß für deren Genauigkeit verwendet werden. Während jedoch zur Bildung des *Spielrisikos* bei einem mathematisch geordneten Spiele entweder nur die Verluste oder nur die Gewinne herangezogen zu werden brauchen (s. Nr. 122), sind bei der Bildung des *Fehlerrisikos* alle möglichen Fehler in Rechnung zu nehmen. Dagegen entsteht die Frage, wonach man die Bedeutung eines Fehlers, den mit ihm verbundènen Nachteil, beurteilen soll. Wird zunächst allgemein die Bedeutung eines Fehlers ε als Funktion seines Wertes aufgefaßt und mit $\psi(\varepsilon)$ bezeichnet, so ergibt

sich für das Fehlerrisiko[1]) die Definition: *es sei die Summe der Produkte aus den Wahrscheinlichkeiten aller möglichen Fehler mit den ihnen entsprechenden Funktionswerten* $\psi(\varepsilon)$. Ist $\varphi(\varepsilon)$ das Fehlergesetz, so drückt sich das Fehlerrisiko R durch die Formel aus:

$$R = \int \psi(\varepsilon)\varphi(\varepsilon)\,d\varepsilon, \tag{1}$$

die Integration über alle Werte von ε ausgedehnt.

Es liegt in der Natur der Sache, dem Fehler eine von seinem Zeichen unabhängige und mit seinem Betrage wachsende Bedeutung beizulegen, d. h. die Funktion $\psi(\varepsilon)$ so zu wählen, daß

$$\psi(\varepsilon) = \psi(-\varepsilon), \text{ und } \psi(\varepsilon') > \psi(\varepsilon), \text{ wenn } |\varepsilon'| > |\varepsilon|. \tag{2}$$

Bleibt man bei diesen allgemeinen Festsetzungen stehen, so läßt sich zunächst nachweisen, daß, sofern der einer experimentellen Bestimmung anhaftende Fehler das Gesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2(\varepsilon - a)^2}$$

befolgt, das mit jener Bestimmung verbundene Fehlerrisiko am *kleinsten* wird, wenn $a = 0$ ist.

Aus

$$R = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} \psi(\varepsilon)\, e^{-h^2(\varepsilon - a)^2}\, d\varepsilon$$

folgt nämlich

$$\frac{dR}{da} = \frac{2h^3}{\sqrt{\pi}} \int_{-\infty}^{\infty} \psi(\varepsilon)(\varepsilon - a)\, e^{-h^2(\varepsilon - a)^2}\, d\varepsilon,$$

$$= \frac{2h^3}{\sqrt{\pi}} \int_{-\infty}^{\infty} \psi(t + a)\, t e^{-h^2 t^2}\, dt;$$

zerlegt man das Integrationsgebiet durch die Null in zwei Teile, ändert in dem ersten Teile des Integrals das Vorzeichen der Variablen t, so ergibt sich vermöge (2) für den Differentialquotienten die Darstellung:

$$\frac{dR}{da} = \frac{2h^3}{\sqrt{\pi}} \int_0^{\infty} [\psi(t + a) - \psi(t - a)]\, t e^{-h^2 t^2}\, dt.$$

Hier nimmt t nur positive Werte an; ist daher $a > 0$, so ist nach (2)

$$\psi(t + a) - \psi(t - a) > 0, \text{ also auch } \frac{dR}{da} > 0;$$

1) Vielfach auch als „totales Fehlerrisiko" bezeichnet.

dagegen bestehen für $a < 0$ die umgekehrten Beziehungen. Es hat also $\frac{dR}{da}$ immer gleiches Vorzeichen mit a, folglich erlangt R für $a = 0$ sein Minimum.

Das Fehlerrisiko wird ferner um so kleiner, je größer h ist, vorausgesetzt, daß $a = 0$ ist. Denn es ist dann

$$R = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} \psi(\varepsilon) e^{-h^2 \varepsilon^2} d\varepsilon = \frac{2}{\sqrt{\pi}} \int_{0}^{\infty} \psi\left(\frac{t}{h}\right) e^{-t^2} dt,$$

folglich das zu einem Werte von t gehörige Element des Integrals und daher R selbst um so kleiner, je größer h ist.

Durch diese Beziehung des Fehlerrisikos zum Präzisionsmaß ist erwiesen, daß jedes Fehlerrisiko, das mit Einhaltung der Relationen (2) gerechnet ist, sich zur *Beurteilung der Genauigkeit* eignet.

Es ist wichtig, zu bemerken, daß das Fehlerrisiko R gleichbedeutend ist mit dem *Mittel-* oder *Durchschnittswerte der Funktion* $\psi(\varepsilon)$ (s. Nr. 55). Läge demnach eine Reihe von Werten $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ des Fehlers vor, wie sie der Zufall zusammengetragen hat, d. h. wie sie sich bei einer Reihe von Beobachtungen ergaben, so darf dem Bernoullischen Theorem zufolge

$$\frac{1}{n}\{\psi(\varepsilon_1) + \psi(\varepsilon_2) + \cdots + \psi(\varepsilon_n)\}$$

um so sicherer als ein Näherungswert von R angesehen werden, je größer n ist; denn mit beständig wachsendem n nähert sich dieser Ausdruck der Grenze R.

141. Der Durchschnittsfehler. Der aus der Wahl

$$\psi(\varepsilon) = |\varepsilon|,$$

die den Bedingungen (2) der vorigen Nummer entspricht, hervorgehende Wert von R wird als *Durchschnittsfehler* bezeichnet. Seine Bedeutung ist die des Mittelwertes der Fehlerbeträge. Liegt demnach eine Reihe von n Fehlern: $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ vor, so ist[1])

$$\frac{[|\varepsilon|]}{n} \tag{1}$$

der aus ihr abgeleitete Näherungswert des Durchschnittsfehlers ϑ.

Befolgen die Fehler das Gesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2},$$

so ist

$$\vartheta = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} |\varepsilon| e^{-h^2 \varepsilon^2} d\varepsilon = \frac{2}{h\sqrt{\pi}} \int_{0}^{\infty} t e^{-t^2} dt,$$

---

1) Die eckige Klammer soll hier und im folgenden als *Summenzeichen* für gleichartige Größen zur Verwendung kommen.

und da $\frac{1}{2}$ der Wert des Integrals,

$$\vartheta = \frac{1}{h\sqrt{\pi}}. \tag{2}$$

Hiernach ist der Durchschnittsfehler dem Präzisionsmaß umgekehrt, nach der in Nr. 137 getroffenen Festsetzung *der Genauigkeit* ebenfalls *umgekehrt* proportional.

Bei bekanntem ϑ gibt die Formel (2) ein Mittel zur Bestimmung des Präzisionsmaßes.

142. Der mittlere Fehler. Die Quadratwurzel aus demjenigen Fehlerrisiko, das aus der Wahl

$$\psi(\varepsilon) = \varepsilon^2$$

hervorgeht, ist von Gauß[1]) unter dem Namen des *mittleren Fehlers* in die Theorie eingeführt worden. Der mittlere Fehler ist hiernach die Quadratwurzel aus dem Mittelwert des Fehlerquadrates und wird aus einer Reihe von Fehlerwerten $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ gefunden, indem man aus dem Quotienten

$$\frac{[\varepsilon\varepsilon]}{n} \tag{1}$$

die Quadratwurzel zieht.

Befolgen die Fehler das Gesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2},$$

so ergibt sich für das Quadrat des mittleren Fehlers μ die Formel:

$$\mu^2 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} \varepsilon^2 e^{-h^2\varepsilon^2} d\varepsilon = \frac{2}{h^2\sqrt{\pi}} \int_0^{\infty} t^2 e^{-t^2} dt = \frac{1}{2h^2},$$

woraus

$$\mu = \frac{1}{h\sqrt{2}} \tag{2}$$

folgt.

Auch der mittlere Fehler ist dem Präzisionsmaß und daher der Genauigkeit umgekehrt proportional, eignet sich also wie der Durchschnittsfehler zu einem Genauigkeitsmaß.

Der mittlere Fehler entspricht dem sogenannten mittleren Risiko bei vom Zufall abhängigen Unternehmungen (s. Nr. 125).

143. Der wahrscheinliche Fehler. Dieser ist nicht auf den Begriff des Fehlerrisikos aufgebaut. Man versteht darunter jene

---

1) Theoria combinationis, Art. 7.

Fehler*grenze* r, für deren Überschreitung wie Nichtüberschreitung die Wahrscheinlichkeit $\frac{1}{2}$ besteht.[1]) Die Gültigkeit des Fehlergesetzes

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2}$$

vorausgesetzt, ergibt sich r aus der Gleichung

$$\frac{1}{2} = \frac{2h}{\sqrt{\pi}} \int_0^r e^{-h^2 \varepsilon^2} d\varepsilon = \frac{2}{\sqrt{\pi}} \int_0^{rh} e^{-t^2} dt,$$

deren Lösung nach Nr. 79, Gleichung (10),

$$rh = 0{,}476936 = \varrho_0 \tag{1}$$

ist. Es eignet sich demnach r, als dem Präzisionsmaß umgekehrt proportional, theoretisch in gleicher Weise als Genauigkeitsmaß wie ϑ und μ.

Liegt eine durch den Zufall erzeugte Reihe von n Fehlern vor, und ist sie nach den absoluten Beträgen steigend geordnet, so bezeichnet bei ungeradem n der mittelste, bei geradem n ein zwischen den beiden mittelsten angenommener Zahlwert eine angenäherte Bestimmung von r.[2])

144. Vergleichende Betrachtung der Genauigkeitsmaße. Mit den drei in den letzten Nummern besprochenen ist die Zahl der *möglichen* Genauigkeitsmaße nicht erschöpft. Aus dem allgemeinen Begriffe des Fehlerrisikos ließe sich vielmehr durch Spezialisierung der Funktion $\psi(\varepsilon)$ noch eine unabsehbare Reihe geeigneter Maße ableiten.

Unter diesen sind diejenigen, welche nach Art des Durchschnitts- und des mittleren Fehlers aus *Mittelwerten der Fehlerpotenzen* hervorgehen, Gegenstand eingehender Untersuchungen geworden. Wählt man nämlich

$$\psi(\varepsilon) = |\varepsilon^m|$$

und bezeichnet den zugehörigen Wert von R mit R_m, so ist bei Geltung des exponentiellen Gesetzes

1) Gauß, Bestimmung der Genauigkeit der Beobachtungen. Zeitschr. f. Astron. I, 1816.

2) Hausdorff (Leipz. Ber., 1901, p. 164—166) empfiehlt den Grundgedanken, der dieser Bestimmung des wahrscheinlichen Fehlers zugrunde liegt, für eine empirische Ermittlungsweise des Präzisionsmaßes h selbst, der der Vorzug einer größeren Sicherheit zukommt (s. d. nächste Nr.): zerlegt man die Reihe der nach ihrer absoluten Größe geordneten Fehler so, daß 84% der Fehler unter und 16% über dem Schnitt liegen, so gibt der reziproke Wert der den Schnitt bewirkenden Fehlergröße eine Bestimmung von h. Vgl. hierzu Helmert, Die Ausgleichungsrechnung etc., 2. Aufl., p. 85 ff.

$$R_m = \frac{2h}{\sqrt{\pi}} \int_0^\infty \varepsilon^m e^{-h^2 \varepsilon^2} d\varepsilon = \frac{2}{h^m \sqrt{\pi}} \int_0^\infty t^m e^{-t^2} dt = \frac{\Gamma\left(\frac{m+1}{2}\right)}{h^m \sqrt{\pi}}, \qquad (1)^{1)}$$

folglich die Größe $\sqrt[m]{R_m}$ dem Präzisionsmaß umgekehrt proportional und daher ebenso als Genauigkeitsmaß verwendbar wie ϑ und μ.

Bei Entscheidung der Frage, welches von all diesen Genauigkeitsmaßen zu wählen sei, hat — den praktischen Gesichtspunkt möglichst einfacher Rechnung zunächst beiseite gelassen — die folgende theoretische Erwängung mitzusprechen. In Wirklichkeit wird für R_m immer nur aus einer vorgelegten Reihe von Einzelwerten des ε: $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$, ein Näherungswert zu erzielen sein durch Ausrechnung der Formel

$$\frac{[\varepsilon^m]}{n}, \text{ beziehungsweise } \frac{[|\varepsilon|^m]}{n}, \qquad (2)$$

je nachdem m gerad oder ungerad ist. Nun erweist eine eingehende Untersuchung[2]), daß der Ausdruck (2) nicht für jedes m gleichrasch dem theoretischen Werte R_m als Grenze sich nähert, mit andern Worten, daß bei gegebenem (großen) n die Grenzen, innerhalb welcher die Differenz

$$R_m - \frac{[\varepsilon^m]}{n}$$

mit gegebener Wahrscheinlichkeit zu erwarten ist, von m abhängen und am engsten sind für $m = 2$, *daß also die Beurteilung der Genauigkeit nach dem Mittelwerte der Fehlerquadrate oder nach dem mittleren Fehler die sicherste ist.*

Die Sicherheit, welche die Wahl $m = 1$ — der Durchschnittsfehler — gewährt, kommt nach jenen Untersuchungen der von $m = 2$ sehr nahe. Trotzdem und trotz der einfacheren Rechnung, welche ϑ gegenüber μ erfordert, ist doch der mittlere Fehler das üblichste Genauigkeitsmaß geworden.

Aus der Zusammenstellung der Formeln für ϑ, μ, r:

$$\vartheta = \frac{1}{h\sqrt{\pi}}, \quad \mu = \frac{1}{h\sqrt{2}}, \quad r = \frac{\varrho_0}{h}$$

---

1) Die Gammafunktion, welche den Zähler des Bruches bildet, hat für

$m = 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad \cdots$

den Wert

$$\Gamma\left(\frac{m+1}{2}\right) = 1, \quad \frac{1}{2}\sqrt{\pi}, \quad 1, \quad \frac{1\cdot 3}{2^2}\sqrt{\pi}, \quad 1\cdot 2, \quad \frac{1\cdot 3\cdot 5}{2^3}\sqrt{\pi}, \quad \cdots$$

2) Vgl. hierzu Helmert, Über die Wahrscheinlichkeit der Potenzsummen etc. Zeitschr. f. Math. u. Phys. 21 (1876); Pizzetti l. c., p. 256 ff.; des Verf.s „Theorie der Beobachtungsfehler", p. 130 ff.

geht übrigens hervor, daß mit einer der drei Größen auch die beiden andern und h bestimmt sind; es ergeben sich nämlich die Beziehungen:

$$\begin{aligned} \mu &= \sqrt{\frac{\pi}{2}}\,\vartheta = 1{,}25331\,\vartheta, \\ r &= \varrho_0\sqrt{2}\,\mu = 0{,}67449\,\mu, \\ r &= \varrho_0\sqrt{\pi}\,\vartheta = 0{,}84533\,\vartheta, \\ h &= \frac{1}{\vartheta\sqrt{\pi}} = \frac{1}{\mu\sqrt{2}} = \frac{\varrho^0}{r}, \\ &= \frac{0{,}56419}{\vartheta} = \frac{0{,}70710}{\mu} = \frac{0{,}47936}{r}. \end{aligned} \tag{3}$$

Hat man zwei oder mehrere der Größen unabhängig voneinander bestimmt, so liefert ihre Prüfung an den vorstehenden Beziehungen eine *generelle Probe* dafür, ob die zugrunde gelegten Fehler nahe genug dem Fehlergesetz entsprechen, aus welchem jene Beziehungen hervorgegangen sind.

Eine *eingehende Probe* hätte sich auf die Verteilung der Fehler nach Vorzeichen und absoluter Größe unter Benützung eines, etwa aus μ, gerechneten Präzisionsmaßes zu erstrecken. Unter n Fehlern ist die Hälfte positiv und die andere Hälfte negativ, ferner sind

$$n\frac{2}{\sqrt{\pi}}\int_0^{ah} e^{-t^2}\,dt = n\Phi(ah)$$

Fehler zwischen $-a$ und $+a$ zu erwarten. Diese Angaben entsprechen der *wahrscheinlichsten* Verteilung (s. Nr. 86). Umfangreiche Untersuchungen, welche in dieser Richtung ausgeführt worden sind, haben das Fehlergesetz

$$\frac{h}{\sqrt{\pi}}e^{-h^2\varepsilon^2}$$

als eine gute approximative Darstellung der Fehlerfrequenz bestätigt.[1])

Häufigen Gebrauches wegen ist nachstehend eine Tabelle mitgeteilt, aus welcher die Wahrscheinlichkeit zu entnehmen ist, daß der einer Beobachtung oder Bestimmung anhaftende Fehler dem Betrage nach zwischen 0 und dem k-fachen Durchschnitts-, mittleren und wahrscheinlichen Fehler gelegen sei. Der Ausdruck für diese Wahrscheinlichkeit ist nach den Formeln (3) bezüglich des Durchschnittsfehlers:

---

1) Näheres hierüber vgl. in Verfassers „Theorie der Beobachtungsfehler“, p. 188—202; C. S. Peirce, On the theory of errors of observations. Coast and Geod. Survey 1870.

$$P = \frac{2}{\sqrt{\pi}} \int_0^{kh\vartheta} e^{-t^2} dt = \Phi\left(\frac{k}{\sqrt{\pi}}\right),$$

bezüglich des mittleren Fehlers:

$$P = \frac{2}{\sqrt{\pi}} \int_0^{kh\mu} e^{-t^2} dt = \Phi\left(\frac{k}{\sqrt{2}}\right),$$

bezüglich des wahrscheinlichen Fehlers:

$$P = \frac{2}{\sqrt{\pi}} \int_0^{khr} e^{-t^2} dt = \Phi(k\varrho_0).$$

Wahrscheinlichkeit P, daß der Fehler einer Bestimmung den k-fachen durchschnittlichen, mittleren und wahrscheinlichen Fehler nicht überschreite.

k	$(0, k\vartheta)$	$(0, k\mu)$	$(0, kr)$
	P	P	P
0,2	0,12696	0,16809	0,10731
0,4	0,25072	0,31100	0,21268
0,6	0,42645	0,45122	0,31430
0,8	0,47640	0,57653	0,41052
1,0	0,57489	0,68260	0,50000
1,2	0,66163	0,76956	0,58171
1,4	0,73610	0,83851	0,65498
1,6	0,79789	0,89028	0,71949
1,8	0,85201	0,92818	0,77528
2,0	0,88933	0,95346	0,82261
2,2	0,92074	0,97114	0,86216
2,4	0,94448	0,98763	0,89450
2,6	0,96184	0,99065	0,92051
2,8	0,97445	0,99489	0,94105
3,0	0,98328	0,99729	0,95698
3,2	0,98932	0,99873	0,96910
3,4	0,99331	0,99932	0,97817
3,6	0,99591	0,99968	0,98482
3,8	0,99755	0,99985	0,98962
4,0	0,99857	0,99994	0,99302

An diese Tabelle soll die Frage nach dem mutmaßlichen *größten Fehler* in einer Beobachtungsreihe angeschlossen werden. Bei einer Gattung von Beobachtungen bekannter Genauigkeit von einem größten Fehler schlechtweg zu sprechen, geht nicht an; denn bis zu welcher Größe Fehler auftreten, wenn man die dem Gesetze entsprechende Verteilung voraussetzt, hängt von dem Umfange der ausgeführten Beobachtungsreihe ab. Ist n die Anzahl der Beobachtungen, P die Wahrscheinlichkeit, daß der absolute Betrag eines Fehlers $k\mu$ nicht überschreite, so ist $n(1 - P)$ die erwartungsmäßige Anzahl der Fehler

jenseits dieser Grenze; und nur wenn diese Anzahl 1 oder darüber ist, sind Fehler zu erwarten, die über den Betrag $k\mu$ hinausgehen. Man kann daher durch Auflösung der Gleichung

$$n(1 - P) = 1$$

nach n dem Umfang der Beobachtungsreihe feststellen, bei welcher $k\mu$ als größter Fehler zu erwarten ist. Mit Benützung der oben zusammengestellten Werte von P ergibt sich:

für $k =$ 2	2,4	2,8	3	3,4	3,8	4
$n =$ 21,	81,	196,	389,	1471,	6667,	16667.

Man ersieht daraus, wie rasch n im Vergleiche zu k wächst, und daß nur bei sehr umfangreichen Beobachtungsreihen ein Fehler zu gewärtigen ist, der den dreifachen mittleren Fehler übersteigt. Selbst bei 100 Beobachtungen wird der größte Fehler voraussichtlich nicht über das $2\frac{1}{2}$ fache des mittleren Fehlers betragen.

145. Verwendung von Beobachtungsdifferenzen zur Genauigkeitsbestimmung. Es liegt nahe, Differenzen von Beobachtungen einer Größe zur Beurteilung der Genauigkeit zu verwenden, weil sie, unabhängig von der Kenntnis des wahren Wertes der beobachteten Größe, Differenzen von Beobachtungsfehlern darstellen. Liegen nämlich für eine unbekannte Größe X zwei Beobachtungsresultate l_1, l_2 vor, so ist

$$X = l_1 + \varepsilon_1 = l_2 + \varepsilon_2,$$

daher

$$\Delta = l_1 - l_2 = \varepsilon_2 - \varepsilon_1.$$

Befolgen die Einzelfehler das Gesetz $\frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2}$, so unterliegen die Differenzen (s. Nr. 138 Schluß) dem Gesetz

$$\frac{h}{\sqrt{2\pi}} e^{-\frac{h^2}{2}\Delta^2};$$

hiernach ist der Mittelwert von $|\Delta|$ gleich $\frac{\sqrt{2}}{h\sqrt{\pi}}$ (s. Nr. 141), also auch gleich $\vartheta\sqrt{2}$, und der Mittelwert von Δ^2 gleich $\frac{1}{h^2}$ (s. Nr. 142), mithin gleich $2\mu^2$. Stehen also s (unabhängige) Beobachtungsdifferenzen $\Delta_1, \Delta_2, \cdots \Delta_s$ zur Verfügung, so kann man, je größer s, um so sicherer setzen:

$$\vartheta\sqrt{2} = \frac{[|\Delta|]}{s}$$

$$2\mu^2 = \frac{[\Delta\Delta]}{s},$$

woraus sich für ϑ und μ die Bestimmungen ergeben:

$$\vartheta = \frac{[|\Delta|]}{s\sqrt{2}}, \tag{1}$$

$$\mu = \sqrt{\frac{[\Delta\Delta]}{2s}}. \tag{2}$$

Aus n Beobachtungswerten $l_1, l_2, \cdots l_n$ von X lassen sich $s = \frac{n(n-1)}{2}$ Beobachtungsdifferenzen bilden, die indessen nicht sämtlich unabhängig voneinander sind; vielmehr lassen sich nur $n-1$ unabhängige auswählen; trotzdem gelten, wie Andrae[1]) und Helmert[2]) nachgewiesen haben, bei Verwendung aller die Formeln:

$$\vartheta = \frac{[|\Delta|]\sqrt{2}}{n(n-1)}, \tag{3}$$

$$\mu = \sqrt{\frac{[\Delta\Delta]}{n(n-1)}}. \tag{4}$$

II. Abschnitt. Kombination von Beobachtungen.

§ 1. Direkte Beobachtungen.

146. Direkte Beobachtungen gleicher Genauigkeit. Das arithmetische Mittel als die mit dem kleinsten Fehlerrisiko verbundene Bestimmung. Man spricht von direkten Beobachtungen, wenn die Messung an der zu bestimmenden Größe selbst vorgenommen wird.

An einer unbekannten physischen Größe X seien n *gleich genaue* Beobachtungen mit den Resultaten $l_1, l_2, \cdots l_n$ vorgenommen worden. Wir nehmen an, es lasse sich im voraus eine Kombinationsregel angeben, durch deren Anwendung sich aus den Beobachtungsergebnissen der vorteilhafteste Wert der Unbekannten x ergibt; sie laute:

$$x = f(l_1, l_2, \cdots l_n). \tag{1}$$

Vorausgesetzt, daß die Fehler der Beobachtungen dem Gesetze

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2}$$

folgen, soll jener Wert als der vorteilhafteste gelten, der das kleinste Fehlerrisiko nach sich zieht.

Dies sind die Grundlagen für die folgende Untersuchung.

---

1) Astron. Nachr. LXXIV (1869). 2) Ibid. LXXXVIII (1876).

Bezeichnet man die Ableitung von f in bezug auf l_i mit α_i, so hat man folgende Entwicklung für $f(X, X, \cdots X)$:

$$f(X, X, \cdots X) = f(l_1 + X - l_1, l_2 + X - l_2, \cdots l_n + X - l_n)$$
$$= x + \alpha_1 (X - l_1) + \alpha_2 (X - l_2) + \cdots + \alpha_n (X - l_n) + \Omega,$$

wobei Ω die Zusammenfassung der Glieder höherer Ordnung in bezug auf die Differenzen $X - l_1$, $X - l_2$, $\cdots$ bezeichnet. Setzt man diese Differenzen als sehr klein voraus und unterdrückt in Rücksicht darauf Ω, so wird weiter:

$$f(X, X, \cdots X) = x + X[\alpha] - \alpha_1 l_1 - \alpha_2 l_2 - \cdots - \alpha_n l_n,$$

und daraus ergibt sich, wenn man sich der Abkürzung

$$f(X, X, \cdots X) - X[\alpha] = k \tag{2}$$

bedient:

$$x = k + \alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n. \tag{3}$$

Damit x durch die Beobachtungsergebnisse allein bestimmt sei, muß k eine von X unabhängige Größe bedeuten.

Schreibt man (2) in der Form

$$f(X, X, \cdots X) = k + \alpha_1 X + \alpha_2 X + \cdots + \alpha_n X, \tag{2*}$$

so folgt aus (2*) und (3) durch Subtraktion:

$$f(X, X, \cdots X) - x = \alpha_1 \varepsilon_1 + \alpha_2 \varepsilon_2 + \cdots + \alpha_n \varepsilon_n.$$

Diese lineare Funktion der Beobachtungsfehler befolgt nach Nr. 138 das Gesetz

$$\frac{H}{\sqrt{\pi}} e^{-H^2 z^2},$$

wobei

$$H^2 = \frac{h^2}{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}$$

ist.

Setzt man also

$$f(X, X, \cdots X) - x = z,$$

so wird

$$X - x = X - f(X, X, \cdots X) + z$$
$$= X - X[\alpha] - k + z$$
$$= X(1 - [\alpha]) - k + z$$

und mit der Abkürzung

$$X(1 - [\alpha]) - k = A$$

weiter

$$X - x = A + z;$$

wird demnach der Fehler in der Bestimmung x, d. i. $X - x$, mit t bezeichnet, so ist vermöge $z = t - A$ und $dz = dt$:

$$\frac{H}{\sqrt{\pi}} e^{-H^2(t-A)^2} dt$$

die Wahrscheinlichkeit, daß die Abweichung des vorteilhaftesten von dem wahren Werte zwischen t und $t + dt$ falle.

Nach den Ergebnissen der Untersuchung in Nr. 140 wird das der Bestimmung von x anhaftende Fehlerrisiko am kleinsten sein, wenn

$$A = 0 \quad \text{und} \quad H \text{ ein Maximum}$$

ist.

Die erste Forderung führt zu der Gleichung

$$X(1 - [\alpha]) - k = 0,$$

die bei der Unabhängigkeit zwischen k und X nur erfüllt ist durch

$$k = 0 \quad \text{und} \quad [\alpha] = 1.$$

Hiernach sind die Koeffizienten α des Ausdrucks

$$x = \alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n, \tag{4}$$

damit er den vorteilhaftesten Wert darstelle, gemäß den Forderungen:

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1, \quad \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 \text{ ein Minimum}$$

zu bestimmen. Dies aber führt auf

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = \frac{1}{n};$$

demnach ist

$$x = \frac{l_1 + l_2 + \cdots + l_n}{n}; \tag{5}$$

d. h. *unter den eingangs gemachten Voraussetzungen ist das arithmetische Mittel aus den Beobachtungsergebnissen die mit dem kleinsten Fehlerrisiko behaftete Bestimmung der Unbekannten.*

Das Gesetz, welchem die Fehler dieser Bestimmung folgen, lautet den vorstehenden Entwicklungen zufolge

$$\frac{h\sqrt{n}}{\sqrt{\pi}} e^{-n h^2 z^2},$$

ihr Präzisionsmaß ist

$$H = h\sqrt{n},$$

ihr mittlerer Fehler

$$\mu_x = \frac{1}{H\sqrt{2}} = \frac{1}{h\sqrt{2n}} = \frac{\mu}{\sqrt{n}}, \tag{6}$$

wenn μ den mittleren Fehler einer Beobachtung bedeutet.

Die Genauigkeit des arithmetischen Mittels aus n gleich genauen Beobachtungen ist demnach $\sqrt{n}$-mal größer als die Genauigkeit der Beobachtungen selbst.

147. Fortsetzung. *Das arithmetische Mittel als wahrscheinlichster Wert der Unbekannten.* Zur Bestimmung einer unbekannten Größe X liegen n gleich genaue Beobachtungen $l_1, l_2, \cdots l_n$ vor, deren wahre Fehler

$$\begin{aligned} \varepsilon_1 &= -l_1 + X \\ \varepsilon_2 &= -l_2 + X \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ \varepsilon_n &= -l_n + X \end{aligned} \tag{1}$$

das Gesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2} \tag{2}$$

befolgen mögen. Dann ist die Wahrscheinlichkeit a priori, daß n erst anzustellende Beobachtungen mit den Fehlern $\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ behaftet sein werden, proportional dem Ausdruck

$$e^{-h^2(\varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2)};$$

liegen aber die Beobachtungen bereits vor, so hat die Annahme, x sei der wahre Wert der beobachteten Größe, eine dem Ausdruck

$$e^{-h^2[(-l_1+x)^2 + (-l_2+x)^2 + \cdots + (l_n + x)^2]}$$

proportionale Wahrscheinlichkeit; diese wird am größten, wenn

$$(-l_1 + x)^2 + (-l_2 + x)^2 + \cdots + (-l_n + x)^2 \text{ ein Minimum} \tag{3}$$

wird. Daraus ergibt sich für x die Bestimmung:

$$x = \frac{l_1 + l_2 + \cdots + l_n}{n}. \tag{4}$$

Das arithmetische Mittel ist also, die Geltung von (2) vorausgesetzt, derjenige Wert der Unbekannten, welcher der wahrscheinlichsten Hypothese entspricht, ihr *wahrscheinlichster Wert.*

In der Bedingung (3) spricht sich jenes Prinzip in seiner einfachsten Form aus, das die Grundlage des Ausgleichungsverfahrens „nach der *Methode der kleinsten Quadrate*“ bildet. Dieses Prinzip lautet dahin, *daß der vorteilhafteste Wert der Unbekannten derjenige sei, für welchen die Summe der Quadrate der den Beobachtungen zugeschriebenen (scheinbaren) Fehler ein Minimum ist.*[1])

---

1) Zur Geschichte der Erfindung und Begründung der Methode der kleinsten Quadrate genüge die folgende kurze Notiz. Die erste Veröffentlichung und auch der Name stammen von A. M. Legendre und sind in dem vom 6. März 1805 datierten, 1806 gedruckten, 9 Seiten umfassenden „Appendice“ zu der Schrift: „Nouvelles méthodes pour la détermination des orbites des comètes“ enthalten. Der Zeitfolge der Publikation nach kommt der Amerikaner R. Adrain mit seiner

148. Bestimmung der Genauigkeit der Beobachtungen und ihres arithmetischen Mittels. Zur Erledigung dieser Aufgabe legen wir uns folgende Frage vor: Wenn die wahren Fehler der Beobachtungen $l_1, l_2, \cdots l_n$, d. i.

$$\begin{aligned} \varepsilon_1 &= -l_1 + X \\ \varepsilon_2 &= -l_2 + X \\ &\cdot \cdot \cdot \cdot \cdot \\ \varepsilon_n &= -l_n + X, \end{aligned} \tag{1}$$

das Gesetz

$$\frac{h}{\sqrt{\pi}} \cdot e^{-h^2 \varepsilon^2} \tag{2}$$

befolgen; welchem Gesetze sind dann ihre Abweichungen vom arithmetischen Mittel, die scheinbaren Fehler

$$\begin{aligned} \lambda_1 &= -l_1 + x \\ \lambda_2 &= -l_2 + x \\ &\cdot \cdot \cdot \cdot \cdot \\ \lambda_n &= -l_n + x, \end{aligned} \tag{3}$$

unterworfen?

Zunächst ergibt sich aus der Summierung von (3) mit Rücksicht auf den Wert von x, daß

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0 \tag{4}$$

ist.

Weiter gibt die Subtraktion homologer Gleichungen aus den Systemen (1) und (3):

$$\begin{aligned} \varepsilon_1 - \lambda_1 &= X - x \\ \varepsilon_2 - \lambda_2 &= X - x \\ &\cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ \varepsilon_n - \lambda_n &= X - x; \end{aligned} \tag{5}$$

1808 im I. Bande der von ihm selbst herausgegebenen Zeitschrift „The analyste" niedergelegten Abhandlung „Research concerning the probabilities of the errors etc.", der wohl unzweifelhaft unabhängig von andern zu der Methode gelangt war. Das Prioritätsrecht der Erfindung gebührt aber C. F. Gauß, der nach mehrfachen eigenen Angaben (s. den 1900 erschienenen Band VIII seiner Werke, p. 136—141) sich seit 1794 im Besitze der Methode befand; er gab ihr auch die erste wissenschaftliche Begründung in der „Theoria motus corporum coelestium", Hamburg 1809. Auf ihn folgte P. S. de Laplace mit seinen tiefgehenden Untersuchungen über den Gegenstand im IV. Kapitel seiner 1812 zum ersten Male erschienenen „Théorie analytique des probabilités". Einen gewissen Abschluß in der theoretischen Entwicklung bedeutet die große 1821—1826 in den Schriften der Göttinger gel. Gesellschaft von C. F. Gauß veröffentlichte Abhandlung: „Theoria combinationis observationum erroribus minimis obnoxiae". — Näheres zur geschichtlichen Seite s. in des Verf.s „Theorie der Beobachtungsfehler".

die Summe dieser Gleichungen ist vermöge (4)

$$\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n = n(X - x). \tag{6}$$

Mit Hilfe von (5) und (6) lassen sich die scheinbaren Fehler durch die wahren ausdrücken wie folgt:

$$\lambda_1 = \frac{n-1}{n}\varepsilon_1 - \frac{1}{n}\varepsilon_2 - \frac{1}{n}\varepsilon_3 - \cdots - \frac{1}{n}\varepsilon_n$$

$$\lambda_2 = -\frac{1}{n}\varepsilon_1 + \frac{n-1}{n}\varepsilon_2 - \frac{1}{n}\varepsilon_3 - \cdots - \frac{1}{n}\varepsilon_n \text{ usw.}$$

Es stellt sich also jeder scheinbare Fehler als lineare Funktion der wahren Fehler dar, befolgt daher auch ein Gesetz von der Form (2), jedoch mit einem andern Präzisionsmaß, das für alle λ dasselbe und nach Gleichung (7), Nr. 138 gleich ist

$$h' = \frac{h}{\sqrt{(n-1)\frac{1}{n^2} + \left(\frac{n-1}{n}\right)^2}} = \sqrt{\frac{n}{n-1}}\,h.$$

Die Abweichungen vom arithmetischen Mittel befolgen also das Gesetz

$$\frac{\sqrt{\frac{n}{n-1}}\,h}{\sqrt{\pi}} e^{-\frac{n}{n-1}h^2\lambda^2}, \tag{7}$$

das sich von (2) durch ein im Verhältnis $\sqrt{n} : \sqrt{n-1}$ größeres Präzisionsmaß unterscheidet.

Der Durchschnittswert von $|\lambda|$ ist also nach Nr. 141

$$\frac{1}{h\sqrt{\pi\frac{n}{n-1}}},$$

findet sich aber aus der vorhandenen Reihe von Einzelwerten $\lambda_1, \lambda_2, \cdots \lambda_n$ annähernd gleich

$$\frac{[|\lambda|]}{n};$$

setzt man beide Ausdrücke einander gleich und beachtet, daß $\frac{1}{h\sqrt{\pi}}$ der *Durchschnittsfehler* ϑ *einer Beobachtung* ist, so ergibt sich für diesen die Formel:

$$\vartheta = \frac{[|\lambda|]}{\sqrt{n(n-1)}}. \tag{8}$$

Mit Hilfe von (7) erhält man ferner für den mittleren Wert von λ^2 nach Nr. 142 den Ausdruck

$$\frac{1}{2\frac{n}{n-1}h^2};$$

aus den zu Gebote stehenden Einzelwerten $\lambda_1, \lambda_2, \cdots \lambda_n$ ergibt er sich annähernd zu

$$\frac{[\lambda\lambda]}{n};$$

setzt man beide Ausdrücke einander gleich und berücksichtigt, daß $\frac{1}{h\sqrt{2}}$ der *mittlere Fehler* μ *einer Beobachtung* ist, so erhält man für diesen die praktisch verwendbare Formel

$$\mu = \sqrt{\frac{[\lambda\lambda]}{n-1}}. \tag{9}$$

Der *wahrscheinliche Fehler* r *einer Beobachtung* läßt sich nun aus μ oder ϑ durch einfache Multiplikation mit einem Zahlenkoeffizienten (s. Nr. 144) ableiten.

Nach dem am Schlusse von Nr. 146 gefundenen Resultate steht der *mittlere Fehler* μ_x *des arithmetischen Mittels* zu dem mittleren Fehler μ einer Beobachtung in der Beziehung, daß $\mu_x = \frac{\mu}{\sqrt{n}}$; nach Formel (9) ist also

$$\mu_x = \sqrt{\frac{[\lambda\lambda]}{n(n-1)}}. \tag{10}$$

149. Direkte Beobachtungen ungleicher Genauigkeit. Begriff des Gewichtes. Eine unbekannte Größe X sei n-mal, jedoch nicht unter gleichen Umständen beobachtet worden, so daß den Beobachtungsergebnissen $l_1, l_2, \cdots l_n$, deren Fehler einzeln dem Exponentialgesetz unterworfen sein mögen, verschiedene Genauigkeit zukommt; die Präzisionsmaße $h_1, h_2, \cdots h_n$ werden als bekannt vorausgesetzt.

Die Durchführung des in Nr. 146 befolgten Gedankenganges führt zu dem Ergebnis, daß

$$x = \alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n$$

den vorteilhaftesten Wert der Unbekannten darstellt, wenn die Koeffizienten den Bedingungen:

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1, \quad \frac{\alpha_1^2}{h_1^2} + \frac{\alpha_2^2}{h_2^2} + \cdots + \frac{\alpha_n^2}{h_n^2} \text{ ein Minimum}$$

gemäß bestimmt werden; diese Bedingungen ergeben aber:

$$\alpha_1 = \frac{h_1^2}{h_1^2 + h_2^2 + \cdots + h_n^2}, \quad \alpha_2 = \frac{h_2^2}{h_1^2 + h_2^2 + \cdots + h_n^2}, \cdots$$

Demnach ist

$$x = \frac{h_1^2 l_1 + h_2^2 l_2 + \cdots + h_n^2 l_n}{h_1^2 + h_2^2 + \cdots + h_n^2} \tag{1}$$

der vorteilhafteste, d. i. der mit dem kleinsten Fehlerrisiko verbundene Wert der Unbekannten. Es ist dies unter den gemachten Voraussetzungen auch der wahrscheinlichste Wert, weil er den Ausdruck

$$e^{-[h_1^2(-l_1+x)^2+h_2^2(-l_2+x)^2+\cdots+h_n^2(-l_n+x)^2]}$$

zu einem Maximum macht; denn für den Wert (1) von x wird

$$h_1^2(-l_1+x)^2+h_2^2(-l_2+x)^2+\cdots+h_n^2(-l_n+x)^2 \tag{2}$$

ein Minimum.

Man kann die einzelnen Beobachtungen statt durch das Präzisionsmaß durch den mittleren Fehler charakterisieren; ist μ_1 der mittlere Fehler von l_1, so ist

$$h_1^2=\frac{1}{2\mu_1^2}$$

zu setzen usw.; dann geht (1) über in

$$x=\frac{\frac{l_1}{\mu_1^2}+\frac{l_2}{\mu_2^2}+\cdots+\frac{l_n}{\mu_n^2}}{\frac{1}{\mu_1^2}+\frac{1}{\mu_2^2}+\cdots+\frac{1}{\mu_n^2}}.$$

Multipliziert man Zähler und Nenner mit der beliebig gewählten positiven Zahl μ^2 und setzt

$$\frac{\mu^2}{\mu_i^2}=p_i \qquad (i=1,2,\cdots n), \tag{3}$$

so wird auch

$$x=\frac{p_1l_1+p_2l_2+\cdots+p_nl_n}{p_1+p_2+\cdots+p_n}=\frac{[pl]}{[p]}. \tag{4}$$

Die positiven Zahlen $p_1, p_2\cdots p_n$, welche dem Gange der Rechnung nach *proportional* sind *den Quadraten der Genauigkeit* der Beobachtungen, bezeichnet man als deren *Gewichte*. Man kann durch entsprechende Wahl von μ^2 bewirken, daß die Gewichte ganze Zahlen werden oder wenigstens bis auf zu vernachlässigende Bruchteile durch ganze Zahlen ersetzt werden können.

Für $\mu_i=\mu$ wird $p_i=1$; es hat also die eingeführte Zahl μ die Bedeutung des mittleren Fehlers einer fingierten Beobachtung, der das Gewicht 1 zukommt; man nennt aus diesem Grunde μ den *mittleren Fehler der Gewichtseinheit.*

Hätte man $p_1+p_2+\cdots+p_n$ Beobachtungen von der Art der fingierten, also vom Gewichte 1, angestellt und hätten p_1 davon das Resultat l_1, p_2 das Resultat $l_2, \cdots p_n$ das Resultat l_n ergeben, so ergäbe sich aus diesen gleich genauen Beobachtungen auch der Mittelwert (4). *Es wiegt demnach eine Beobachtung, deren Gewicht p ist, p Beobachtungen vom Gewichte 1 auf,* zunächst in bezug auf die Bildung des Mittelwertes.

Das Gewicht tritt zu den bisher betrachteten Präzisionsmaßen als neues hinzu.

Sowie die Form (1) des Resultates aus der Bedingung (2), so kann die Form (4) aus der Bedingung:

$$p_1(-l_1+x)^2+p_2(-l_2+x)^2+\cdots+p_n(-l_n+x)^2 \text{ ein Minimum,} \quad (2^*)$$

abgeleitet werden.

Im Falle ungleich genauer Beobachtungen erweitert sich also das am Schlusse von Nr. 147 erkannte Prinzip dahin, daß *die Summe der mit den Gewichten multiplizierten Quadrate der (scheinbaren) Fehler zu einem Minimum zu machen* sei, um den vorteilhaftesten Wert der Unbekannten zu erhalten.

Die Formel (4) kommt auch zur Anwendung, wenn die $l_1, l_2, \cdots l_n$ nicht unmittelbare Beobachtungsresultate, sondern arithmetische Mittel aus Beobachtungsreihen darstellen; $\mu_1, \mu_2, \cdots \mu_n$ bedeuten dann die mittleren Fehler dieser arithmetischen Mittel. Sind letztere aus gleich genauen Beobachtungen vom mittleren Fehler μ, in den Anzahlen $\nu_1, \nu_2, \cdots \nu_n$, entstanden, so ist

$$\mu_i^2=\frac{\mu^2}{\nu_i};$$

wählt man also die einzelne Beobachtung als Gewichtseinheit, so wird das Gewicht der „Beobachtung" l_i:

$$p_i=\frac{\mu^2}{\mu_i^2}=\nu_i.$$

Für die wahren und scheinbaren Fehler der Beobachtungen bestehen die Gleichungen:

$$\begin{aligned} \varepsilon_1 &= -l_1+X \\ \varepsilon_2 &= -l_2+X \\ &\cdots\cdots \\ \varepsilon_n &= -l_n+X \end{aligned} \qquad (5) \qquad\qquad \begin{aligned} \lambda_1 &= -l_1+x \\ \lambda_2 &= -l_2+x \\ &\cdots\cdots \\ \lambda_n &= -l_n+x. \end{aligned} \qquad (6)$$

Aus dem zweiten System folgt mit Rücksicht auf (4):

$$p_1\lambda_1+p_2\lambda_2+\cdots+p_n\lambda_n=[p\lambda]=0. \qquad (7)$$

Durch Verbindung beider Systeme ergibt sich wegen (7):

$$[p\varepsilon]=[p](X-x),$$

woraus

$$X-x=\frac{p_1}{[p]}\varepsilon_1+\frac{p_2}{[p]}\varepsilon_2+\cdots+\frac{p_n}{[p]}\varepsilon_n;$$

demnach befolgt der Fehler in der Bestimmung x das Gesetz:

$$\frac{H}{\sqrt{\pi}}e^{-H^2z^2},$$

wobei

$$\frac{1}{H^2} = \frac{\frac{p_1^2}{[p]^2}}{h_1^2} + \frac{\frac{p_2^2}{[p]^2}}{h_2^2} + \cdots + \frac{\frac{p_n^2}{[p]^2}}{h_n^2};$$

nun ist aber $h_i^2 = \frac{1}{2\mu_i^2} = \frac{p_i}{2\mu^2}$, folglich

$$H^2 = \frac{[p]}{2\mu^2};$$

daraus aber ergibt sich der *mittlere Fehler des Mittels* x:

$$\mu_x = \frac{1}{H\sqrt{2}} = \frac{\mu}{\sqrt{[p]}}. \tag{8}$$

Mit der Formel (6) in Nr. 146 verglichen zeigt dies, daß das gegenwärtige Mittel äquivalent ist einem gewöhnlichen arithmetischen Mittel aus $[p]$ Beobachtungen vom Gewichte 1.

Setzt man in (8) für die einzelnen p ihre Ausdrücke aus (3), so kommt man zu einer Formel, welche den mittleren Fehler von x durch die mittleren Fehler der zu seiner Berechnung verwendeten l ausdrückt, nämlich:

$$\mu_x = \frac{1}{\sqrt{\frac{1}{\mu_1^2} + \frac{1}{\mu_2^2} + \cdots + \frac{1}{\mu_n^2}}} = \frac{1}{\sqrt{\left[\frac{1}{\mu\mu}\right]}}. \tag{9}$$

Noch handelt es sich darum, mit Hilfe der bekannten Gewichte und der scheinbaren Fehler die Genauigkeitsbestimmung durchzuführen, nämlich den mittleren Fehler der Gewichtseinheit und des Mittels x zu berechnen.

Zwischen dem mittleren Fehler μ_i einer Beobachtung vom Gewichte p_i und dem mittleren Fehler μ einer Beobachtung vom Gewichte 1 besteht nach (3) die Beziehung:

$$\mu = \mu_i \sqrt{p_i}.$$

Überträgt man diese Beziehung auf die einzelnen Abweichungen vom Mittel, so wird der Abweichung λ_i einer Beobachtung vom Gewichte p_i bei einer Beobachtung vom Gewichte 1 die Abweichung

$$\lambda_i' = \lambda_i \sqrt{p_i}$$

entsprechen. Aus den transformierten Abweichungen $\lambda_1', \lambda_2', \cdots \lambda_n'$ aber bestimmt sich der *mittlere Fehler der Gewichtseinheit* gemäß der Formel (9) der vorigen Nummer

$$\mu = \sqrt{\frac{[\lambda'\lambda']}{n-1}};$$

führt man die scheinbaren Fehler λ [nach (6)] selbst ein, so wird also endgültig

$$\mu = \sqrt{\frac{[p\lambda\lambda]}{n-1}}. \tag{10}$$

Mittels derselben Schlüsse erhält man für den *durchschnittlichen Fehler der Gewichtseinheit* die Formel:

$$\vartheta = \frac{[|\sqrt{p}\lambda|]}{\sqrt{n(n-1)}}. \tag{11}$$

Mit Benützung von (8) ergibt sich nun für den *mittleren Fehler des Mittels* der Ausdruck:

$$\mu_x = \sqrt{\frac{[p\lambda\lambda]}{(n-1)[p]}}. \tag{12}$$

Die Formeln (9) und (12) werden im allgemeinen nicht übereinstimmende Werte liefern, weil die erste mit den *mittleren* Fehlern, die zweite mit den *wirklichen* Abweichungen der Beobachtungen rechnet.

150. Zusammenstellung der Resultate. a) Beobachtungen gleicher Genauigkeit.

Man rechnet aus $l_1, l_2, \cdots l_n$:

$$x = \frac{[l]}{n},$$

daraus die

$$\lambda_i = -l_i + x \qquad (i = 1, 2, \cdots n),$$

deren richtige Berechnung mittels der Gleichung $[\lambda] = 0$ kontrolliert werden kann; weiter

$$\mu = \sqrt{\frac{[\lambda\lambda]}{n-1}}$$

und hieraus

$$\mu_x = \frac{\mu}{\sqrt{n}};$$

das Resultat der Ausgleichung stellt man kurz durch

$$x \mp \mu_x$$

dar.

Die Summe $[\lambda\lambda]$ rechnet man aus den einzelnen λ mit Hilfe von Quadrattafeln. Man kann sie zur Kontrolle auch aus den l rechnen auf Grund der Gleichung:

$$[\lambda\lambda] = [ll] - \frac{[l]^2}{n},$$

die sich leicht aus den ersten beiden Ansätzen ableiten läßt.

b) Beobachtungen ungleicher Genauigkeit. Sind außer $l_1, l_2, \cdots l_n$ deren Gewichte $p_1, p_2, \cdots p_n$ gegeben, so rechne man

$$x = \frac{[pl]}{[p]},$$

daraus die

$$\lambda_i = -l_i + x,$$

welche bei richtiger Ausführung der Rechnung die Kontrolle $[p\lambda] = 0$ bestehen müssen; dann

$$\mu = \sqrt{\frac{[p\lambda\lambda]}{n-1}}$$

und

$$\mu_x = \frac{\mu}{\sqrt{[p]}};$$

das Resultat ist wieder in der Form

$$x \mp \mu_x$$

zu geben.

Waren die mittleren Fehler der l gegeben, so hat die Gewichtsberechnung voranzugehen; man wählt für die p_i in passender Weise Zahlen, die den $\frac{1}{\mu_i^2}$ proportional sind.

Die Summe $[p\lambda\lambda]$ kann außer aus den λ auch aus den l nach der Formel

$$[p\lambda\lambda] = [pll] - \frac{[pl]^2}{[p]}$$

berechnet werden, die sich aus den zwei ersten Ansätzen ergibt.

Die Zusammenstellung nimmt lediglich auf die Genauigkeitsbestimmung durch den mittleren Fehler Rücksicht. Für weitergehende Untersuchungen sind die nötigen Formeln den vorangehenden Nummern zu entnehmen.

151. Beispiel LV. *Experimentelle Bestimmung der konstanten Wahrscheinlichkeit eines Ereignisses.*

Wir stellen dieses Beispiel aus mehrfachen Gründen an die Spitze Einmal handelt es sich um Beobachtungen, deren Fehler im strengen Sinne den Charakter des Zufälligen besitzen. Ferner ist die unbekannte Größe nicht eine physische, sondern eine Zahlengröße und als solche streng definiert. Schließlich sind Untersuchungen von der hier durchgeführten Art grundlegend für die Behandlung von Fragen der theoretischen Statistik.

a) Aus einer Urne mit sechs mit den Nummern 1 bis 6 bezeichneten Kügelchen sind Ziehungen je eines Kügelchens mit Zurücklegung des jeweilen gezogenen ausgeführt worden, und zwar in

$n = 37$ Reihen zu je $s = 100$ Ziehungen; die gezogenen Nummern wurden notiert.[1])

Das Ereignis E sei das *Erscheinen einer ungeraden Nummer.* Stellt man sich auf den Standpunkt, die Zusammensetzung des Urneninhaltes sei unbekannt, so liefert jede Reihe in dem Quotienten $\frac{m}{100}$, dessen Zähler die Anzahl der erschienenen ungeraden Nummern bedeutet, eine Beobachtung l der unbekannten Wahrscheinlichkeit X für das Erscheinen einer solchen Nummer, und die 37 so erhaltenen Beobachtungen sind als gleich genau zu betrachten, weil sie aus Ziehungsreihen gleichen Umfangs hervorgegangen sind (s. Nr. 85).

Da im vorliegenden Falle der wahre Wert X bekannt, nämlich $\frac{1}{2}$ ist, so ist es möglich, auch die wahren Fehler zu bestimmen und die theoretischen Resultate an dem Beobachtungsmaterial in mannigfacher Richtung zu prüfen.

Das Beobachtungsmaterial sowie die zu dieser Prüfung erforderlichen Zahlenreihen sind in der folgenden Tabelle zusammengestellt.

Nach Addition der ersten Kolonne berechnet man das arithmetische Mittel

$$x = \frac{18{,}54}{37} = 0{,}501,$$

mit Hilfe desselben die λ; davon fallen 21 positiv, 16 negativ aus (gegenüber der wahrscheinlichsten Verteilung 19, 18 oder 18, 19); ihre Summe ist $-0{,}003$ statt 0, was von der Abkürzung des Wertes von x herrührt.

Mit dem wahren Werte $X = \frac{1}{2}$ sind die wahren Fehler ε gerechnet worden.

Für den mittleren Fehler einer Beobachtung ergeben sich die beiden Bestimmungen:

$$\mu = \sqrt{\frac{0{,}067757}{36}} = 0{,}0434,$$

$$\mu = \sqrt{\frac{0{,}0678}{37}} = 0{,}0428,$$

aus diesen das Präzisionsmaß:

$$h = \frac{1}{0{,}0434\sqrt{2}} = 16{,}32,$$

$$h = \frac{1}{0{,}0428\sqrt{2}} = 16{,}53.$$

---

1) Die Versuche sind von meinen Hörern im Wintersemester 1900 ausgeführt worden.

Nr.	Beobachtung l	Scheinb. Fehler λ	Wahre Fehler ε	$\lambda\lambda$	$\varepsilon\varepsilon$	λ, geordnet n. d. absoluten Beträge	ε, geordnet n. d. absoluten Beträge
1	0,56	− 0,059	− 0,06	0,003481	0,0036	0,001	0,00
2	0,47	+ 0,031	+ 0,03	0,000961	0,0009	0,001	0,00
3	0,49	+ 0,011	+ 0,01	0,000121	0,0001	0,001	0,00
4	0,46	+ 0,041	+ 0,04	0,001681	0,0016	0,001	0,00
5	0,52	− 0,019	− 0,02	0,000361	0,0004	0,009	0,01
6	0,53	− 0,029	− 0,03	0,000841	0,0009	0,009	0,01
7	0,58	− 0,079	− 0,08	0,006241	0,0064	0,011	0,01
8	0,46	+ 0,041	+ 0,04	0,001681	0,0016	0,011	0,01
9	0,50	+ 0,001	0,00	0,000001	0,0000	0,011	0,01
10	0,42	+ 0,081	+ 0,08	0,006561	0,0064	0,011	0,01
11	0,54	− 0,039	− 0,04	0,001521	0,0016	0,019	0,02
12	0,52	− 0,019	− 0,02	0,000361	0,0004	0,019	0,02
13	0,56	− 0,059	− 0,06	0,003481	0,0036	0,019	0,02
14	0,50	+ 0,001	0,00	0,000001	0,0000	0,019	0,02
15	0,44	+ 0,061	+ 0,06	0,003721	0,0036	0,021	0,02
16	0,49	+ 0,011	+ 0,01	0,000121	0,0001	0,029	0,03
17	0,56	− 0,059	− 0,06	0,003481	0,0036	0,031	0,03
18	0,50	+ 0,001	0,00	0,000001	0,0000	0,031	0,03
19	0,56	− 0,059	− 0,06	0,003481	0,0036	0,031	0,03
20	0,47	+ 0,031	+ 0,03	0,000961	0,0009	0,031	0,03
21	0,47	+ 0,031	+ 0,03	0,000961	0,0009	0,039	0,04
22	0,51	− 0,009	− 0,01	0,000081	0,0001	0,039	0,04
23	0,52	− 0,019	− 0,02	0,000361	0,0004	0,041	0,04
24	0,48	+ 0,021	+ 0,02	0,000441	0,0004	0,041	0,04
25	0,56	− 0,059	− 0,06	0,003481	0,0036	0,051	0,05
26	0,43	+ 0,071	+ 0,07	0,005041	0,0049	0,051	0,05
27	0,58	− 0,079	− 0,08	0,006241	0,0064	0,059	0,06
28	0,52	− 0,019	− 0,02	0,000361	0,0004	0,059	0,06
29	0,49	+ 0,011	+ 0,01	0,000121	0,0001	0,059	0,06
30	0,49	+ 0,011	+ 0,01	0,000121	0,0001	0,059	0,06
31	0,44	+ 0,061	+ 0,06	0,003721	0,0036	0,059	0,06
32	0,50	+ 0,001	0,00	0,000001	0,0000	0,061	0,06
33	0,54	− 0,039	− 0,04	0,001521	0,0016	0,061	0,06
34	0,51	− 0,009	− 0,01	0,000081	0,0001	0,071	0,07
35	0,45	+ 0,051	+ 0,05	0,002601	0,0025	0,079	0,08
36	0,47	+ 0,031	+ 0,03	0,000961	0,0009	0,079	0,08
37	0,45	+ 0,051	+ 0,05	0,002601	0,0025	0,081	0,08
	18,54 [l]	+ 0,651 − 0,654	+ 0,63 − 0,67	0,067757 [$\lambda\lambda$]	0,0678 [$\varepsilon\varepsilon$]	1,305 [$\vert\lambda\vert$]	1,30 [$\vert\varepsilon\vert$]
		− 0,003 [λ]	− 0,04 [ε]				

Ebenso erhält man für den durchschnittlichen Fehler zwei Bestimmungen:

$$\vartheta = \frac{1,305}{\sqrt{37 \cdot 36}} = 0,0358,$$

$$\vartheta = \frac{1,30}{37} = 0,0351$$

und aus diesen für das Präzisionsmaß die Werte:

$$h = \frac{1}{0{,}0358\sqrt{\pi}} = 15{,}75,$$

$$h = \frac{1}{0{,}0351\sqrt{\pi}} = 16{,}05.$$

Diesen aus den Abweichungen berechneten Werten des Präzisionsmaßes steht derjenige Wert gegenüber, der sich aus der Theorie selbst ergibt (s. Nr. 85), nämlich

$$h = \sqrt{\frac{100}{2 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = 14{,}14.$$

Aus dem ersten Werte von μ berechnet sich der wahrscheinliche Fehler einer Beobachtung:

$$r = 0{,}4769 \cdot 0{,}0434\sqrt{2} = 0{,}0292,$$

durch Abzählung an den in steigender Größe geordneten λ (wie ε) findet man

$$r = 0{,}031, \text{ beziehungsweise } = 0{,}03.$$

Der mittlere Fehler des arithmetischen Mittels, aus dem ersten μ gerechnet, ist

$$\mu_x = \frac{0{,}0434}{\sqrt{37}} = 0{,}0072;$$

der wahre Wert 0,5 liegt innerhalb der hierdurch bestimmten Grenzen:

$$0{,}501 \mp 0{,}007.$$

Die Verteilung der λ nach ihrer absoluten Größe, verglichen mit derjenigen, welche sich aus dem Fehlergesetz

$$\frac{16{,}32}{\sqrt{\pi}} e^{-16{,}32^2 \lambda^2}$$

ergibt, zeigt sich wie folgt:

$\vert\lambda\vert$ zwischen	Anzahl	
	beobachtet	berechnet
0,000 und 0,020	14	13,2
„ „ 0,040	22	23,8
„ „ 0,060	31	30,8
„ „ 0,080	36	34,3
„ „ 0,100	37	36.2

Man kann also sagen, daß die Beobachtungsreihe, obwohl nur von mäßigem Umfange, die theoretischen Resultate in befriedigender Weise wiedergibt.

b) Das Ereignis E sei das *Erscheinen der Nummer 1.* Aus theoretischen Gründen ist hier eine genauere Beobachtungsreihe zu

erwarten; denn der wahre Wert X der Wahrscheinlichkeit von E, $\frac{1}{6}$, gibt mit der entgegengesetzten Wahrscheinlichkeit $\frac{5}{6}$ ein kleineres Produkt als im Falle a), wo beide Wahrscheinlichkeiten $\frac{1}{2}$ waren, und daraus entspringt eine größere Präzision. Hauptsächlich um zu zeigen, daß die Erfahrung diese theoretische Erwartung bestätigt, ist die betreffende Beobachtungsreihe hier mitgeteilt; die Rechnung ist jedoch nur mit den λ geführt.

Nr.	Beobachtung l	Scheinbarer Fehler λ	$\lambda\lambda$	λ, geordnet n. d. absoluten Betrage
1	0,17	— 0,0049	0,000024	0,0049
2	0,20	— 0,0349	0,001218	0,0049
3	0,10	+ 0,0651	0,004238	0,0049
4	0,17	— 0,0049	0,000024	0,0049
5	0,12	+ 0,0451	0,002034	0,0049
6	0,15	+ 0,0151	0,000228	0,0049
7	0,19	— 0,0249	0,000620	0,0051
8	0,22	— 0,0549	0,003014	0,0051
9	0,17	— 0,0049	0,000024	0,0051
10	0,19	— 0,0249	0,000620	0,0149
11	0,14	+ 0,0251	0,000630	0,0149
12	0,22	— 0,0549	0,003014	0,0149
13	0,18	— 0,0149	0,000222	0,0149
14	0,17	— 0,0049	0,000024	0,0151
15	0,13	+ 0,0351	0,001232	0,0151
16	0,12	+ 0,0451	0,002034	0,0151
17	0,18	— 0,0149	0,000222	0,0151
18	0,15	+ 0,0151	0,000228	0,0151
19	0,17	— 0,0049	0,000024	0,0151
20	0,16	+ 0,0051	0,000026	0,0249
21	0,15	+ 0,0151	0,000228	0,0249
22	0,15	+ 0,0151	0,000228	0,0249
23	0,18	— 0,0149	0,000222	0,0251
24	0,16	+ 0,0051	0,000026	0,0251
25	0,14	+ 0,0251	0,000630	0,0251
26	0,18	— 0,0149	0,000222	0,0251
27	0,17	— 0,0049	0,000024	0,0251
28	0,14	+ 0,0251	0,000630	0,0349
29	0,13	+ 0,0351	0,001232	0,0351
30	0,15	+ 0,0151	0,000228	0,0351
31	0,22	— 0,0549	0,003014	0,0451
32	0,19	— 0,0249	0,000620	0,0451
33	0,16	+ 0,0051	0,000026	0,0549
34	0,26	— 0,0949	0,009006	0,0549
35	0,14	+ 0,0251	0,000630	0,0549
36	0,15	+ 0,0151	0,000228	0,0651
37	0,14	+ 0,0251	0,000630	0,0949
	6,11 [l]	+ 0,4569 — 0,4582	0,037524 [$\lambda\lambda$]	0,9151 [λ]
		— 0,0013 [λ]		

Daraus berechnet sich das arithmetische Mittel:

$$x = \frac{6{,}11}{37} = 0{,}1651$$

gegenüber dem wahren Werte 0,1666⋯; der mittlere Fehler einer Beobachtung:

$$\mu = \sqrt{\frac{0{,}037524}{36}} = 0{,}0323$$

das Präzisionsmaß:

$$h = \frac{1}{0{,}0323\sqrt{2}} = 21{,}46,$$

für das sich nach der theoretischen Formel (Nr. 85) der Wert $\sqrt{\frac{100}{2 \cdot \frac{1}{6} \cdot \frac{5}{6}}} = 18{,}97$ ergibt; der durchschnittliche Fehler einer Beobachtung:

$$\vartheta = \frac{0{,}9151}{\sqrt{37 \cdot 36}} = 0{,}0251;$$

das Verhältnis:

$$\frac{\mu}{\vartheta} = 1{,}287,$$

dessen theoretischer Wert (Nr. 144) $\sqrt{\frac{\pi}{2}} = 1{,}236$ ist; der mittlere Fehler des arithmetischen Mittels:

$$\mu_x = \frac{0{,}0323}{\sqrt{37}} = 0{,}0053,$$

so daß der wahre Wert innerhalb der durch x und μ_x bestimmten Grenzen

$$0{,}1651 \mp 0{,}0053$$

enthalten ist. Die beobachtete und die theoretische Verteilung der λ gestaltet sich wie folgt:

$\|\lambda\|$ zwischen	Anzahl	
	beobachtet	berechnet
0,00 und 0,01	9	8,7
„ „ 0,02	19	16,9
„ „ 0,03	27	23,5
„ „ 0,04	30	28,6
„ „ 0,05	32	32,2
„ „ 0,08	36	36,4

Theoretisch war die Genauigkeit dieser zweiten Beobachtungsreihe

$$\sqrt{\frac{\frac{1}{4}}{\frac{5}{36}}} = 1,341$$

mal größer zu erwarten als die der ersten; nach den λ beurteilt, stellt sie sich

$$21,46 : 16,32 = 1,315$$

mal größer dar.

152. Beispiel LVI. Im Laufe der Jahre 1892—1894 wiederholte Bestimmungen der Polhöhe von Kapstadt haben die aus der nachfolgenden Tabelle ersichtlichen 15 Werte ergeben, welche als gleich genaue Beobachtungen zu behandeln und auszugleichen sind.[1])

Nr.	l	λ	$\lambda\lambda$
1	— 33° 56′ 3″,48	— 0,22	0,0484
2	3 ,50	— 0,24	0,0576
3	3 ,50	— 0,24	0,0576
4	3 ,32	— 0,06	0,0036
5	3 ,09	+ 0,17	0,0289
6	2 ,98	+ 0,28	0,0784
7	3 ,07	+ 0,19	0,0361
8	3 ,28	— 0,02	0,0004
9	3 ,27	— 0,01	0,0001
10	3 ,20	+ 0,06	0,0036
11	3 ,30	— 0,04	0,0016
12	3 ,25	+ 0,01	0,0001
13	3 ,11	+ 0,15	0,0225
14	3 ,30	— 0,04	0,0016
15	3 ,27	— 0,01	0,0001
	48″,92	+ 0,86	0,3406
		— 0,88	
		— 0,02	

Man rechnet:

$$x = \frac{48'',92}{15} = 3'',26$$

$$\mu = \sqrt{\frac{0,3406}{14}} = 0'',16$$

$$\mu_x = \frac{0'',16}{\sqrt{15}} = 0'',04$$

1) Bericht über den Stand der Erforschung der Breitenvariation im Dezember 1897, von Th. Albrecht, Berlin 1898, p. 20. — Die l sind hier nicht einfache Beobachtungen, sondern Mittelwerte von Beobachtungsreihen und als solche nicht von völlig gleicher Genauigkeit. Sie sind dem Zwecke entsprechend, dem sie zu dienen hatten, nach der Zeitfolge geordnet.

und kann als Ergebnis der Ausgleichung hinstellen:

$$-33^{0}\,56'\,3'',26 \mp 0'',04.$$

153. Beispiel LVII. Zur Bestimmung der Polhöhe der großen Kuppel des Astrophysikalischen Observatoriums bei Potsdam wurden in den Jahren 1892—1893 wiederholte Beobachtungen nach der Horrebow-Methode an Sternpaaren vorgenommen. Dieselben sind gruppenweise zu Mittelwerten vereinigt worden. Die Mittelwerte der so gebildeten 11 Gruppen nebst den Anzahlen der dabei verwendeten Sternpaare sind in der folgenden Tabelle zusammengestellt. Aus denselben ist das Schlußresultat zu ziehen und seine Genauigkeit zu bestimmen.[1])

Die Anzahlen der Sternpaare werden als Gewichte zu verwenden sein (s. Nr. 149).

Nr.	l	Paare p	pl	λ	$\lambda\lambda$	$p\lambda\lambda$
1	52° 22′ 53″,98	99	97,02	— 0,02	0,0004	0,0396
2	53 ,86	118	101,48	+ 0,10	0,0100	1,1800
3	53 ,95	114	108,30	+ 0,01	0,0001	0,0144
4	54 ,01	142	143,42	— 0,05	0,0025	0,3550
5	54 ,06	163	172,78	— 0,10	0,0100	1,6300
6	53 ,98	192	188,16	— 0,02	0,0004	0,0768
7	53 ,91	158	143,78	+ 0,05	0,0025	0,3950
8	53 ,97	102	98,94	— 0,01	0,0001	0,0102
9	53 ,93	103	95,79	+ 0,03	0,0009	0,0927
10	53 ,90	131	117,90	+ 0,06	0,0036	0,4716
11	53 ,91	90	81,90	+ 0,05	0,0025	0,2250
		1412 [p]	1349,47 [pl]			4,3873 [$p\lambda\lambda$]

Die Kolonne pl ist mit den Resten der l gerechnet, welche nach Abtrennung von 52°22′53″ übrig bleiben.

Man bestimmt nun: das arithmetische Mittel:

$$x = \frac{1349,47}{1412} = 0'',96;$$

den mittleren Fehler der Gewichtseinheit, d. i. der Polhöhenbestimmung aus einem Sternpaare:

$$\mu = \sqrt{\frac{4,3873}{10}} = 0'',66;$$

den mittleren Fehler des Mittels:

$$\mu_x = \frac{0,66}{\sqrt{1412}} = 0'',018.$$

1) Die Polhöhe von Potsdam. 1. Heft. Veröffentlichung des Kgl. Preuß. Geodät. Institutes, Berlin 1898, p. 122.

Das Ergebnis der Ausgleichung ist hiernach:

$$52^0 22' 53'',96 \mp 0'',018.$$

154. Beispiel LVIII. Im Jahre 1880 ist in der Nähe von Berlin eine Grundlinie in zehn nahezu gleichen Teilstrecken von der durchschnittlichen Länge von 234 m gemessen worden. Jede Teilstrecke wurde zweimal, hin und zurück, gemessen. In der nachstehenden Tabelle sind die arithmetischen Mittel der beiden Messungsergebnisse und ihre Differenz angegeben.[1]) Aus den Differenzen (s. Nr. 145) ist zunächst die Genauigkeit der Messung einer Teilstrecke abzuleiten.

Strecke	Mittel aus beiden Messungen in m	Differenz Δ in mm	$\Delta\Delta$
1	236,459 785	— 0,040	0,001 600
2	236,506 768	— 0,026	0,000 676
3	236,492 006	+ 0,240	0,057 600
4	236,471 427	— 0,031	0,000 961
5	243,428 614	+ 1,573	2,474 329
6	232,548 466	+ 0,102	0,010 404
7	228,625 791	— 0,096	0,009 216
8	228,616 588	— 0,120	0,014 400
9	228,552 782	— 1,363	1,857 769
10	228,710 183	— 0,195	0,038 025
	2336,412 310		4,464 980 [$\Delta\Delta$]

Nach Formel (2), Nr. 145, berechnet sich

$$\mu = \sqrt{\frac{4,464980}{20}} = 0^{mm},473$$

als mittlerer Fehler der einmaligen Messung einer Länge von 234 m, so daß der relative mittlere Fehler einer solchen Messung

$$\frac{0,473}{234000} \text{ oder } \frac{1}{500000} \text{ der Länge}$$

ausmacht. Der mittlere Fehler des Mittels aus zwei Messungen einer solchen Strecke bestimmt sich hieraus zu

$$\mu_1 = \frac{0,473}{\sqrt{2}} = 0^{mm},334,$$

beträgt also $\frac{1}{700000}$ der Länge.

155. Funktionen direkt beobachteter Größen. Es sei $F = F(X_1, X_2, \ldots X_n)$ eine beliebige Funktion der unabhängigen

1) Die Neumessung der Grundlinien bei Strehlen, Berlin und Bonn. Veröffentlichung des Kgl. Preuß. Geodät. Institutes, 1897, p. 84ff.

Größen $X_1, X_2, \ldots X_n$; für diese seien durch direkte Beobachtung die vorteilhaftesten Werte $x_1, x_2, \ldots x_n$ mit ihren mittleren Fehlern $\mu_1, \mu_2, \ldots \mu_n$ gefunden worden. Man soll den vorteilhaftesten Wert f von F und seine Genauigkeit, ausgedrückt durch den mittleren Fehler, bestimmen.

Die erste dieser Aufgaben ist unmittelbar gelöst; denn die vorteilhafteste Bestimmung von F ergibt sich aus den vorteilhaftesten Werten der Elemente; also ist

$$f = F(x_1, x_2, \ldots x_n). \tag{1}$$

Für die Lösung der zweiten Aufgabe ist die Grundlage in Nr. 138 gegeben. Sind die wahren Fehler $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n$ der Bestimmungen $x_1, x_2, \ldots x_n$ sehr klein, so kann mit Außerachtlassung von Gliedern mit ihren Potenzen und Produkten gesetzt werden:

$$F = F(x_1 + \varepsilon_1, x_2 + \varepsilon_2, \cdots x_n + \varepsilon_n) = f + \frac{\partial f}{\partial x_1}\varepsilon_1 + \frac{\partial f}{\partial x_2}\varepsilon_2 + \cdots + \frac{\partial f}{\partial x_n}\varepsilon_n,$$

woraus

$$F - f = \frac{\partial f}{\partial x_1}\varepsilon_1 + \frac{\partial f}{\partial x_2}\varepsilon_2 + \cdots + \frac{\partial f}{\partial x_n}\varepsilon_n.$$

Befolgen nun die Fehler $\varepsilon_i \ (i = 1, 2, \ldots n)$ das Gesetz

$$\frac{h_i}{\sqrt{\pi}} e^{-h_i^2 \varepsilon_i^2},$$

so befolgt der Fehler in der Bestimmung f von F das Gesetz

$$\frac{H}{\sqrt{\pi}} e^{-H^2 \varepsilon^2},$$

wobei

$$\frac{1}{H^2} = \frac{\left(\frac{\partial f}{\partial x_1}\right)^2}{h_1^2} + \frac{\left(\frac{\partial f}{\partial x_2}\right)^2}{h_2^2} + \cdots + \frac{\left(\frac{\partial f}{\partial x_n}\right)^2}{h_n^2};$$

nun ist aber

$$\frac{1}{2h_i^2} = \mu_i^2$$

das Quadrat des mittleren Fehlers von x_i,

$$\frac{1}{2H^2} = \mu_f^2$$

das Quadrat des mittleren Fehlers in der Bestimmung f von F; die obige Gleichung ergibt also:

$$\mu_f = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 \mu_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \mu_2^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \mu_n^2}. \tag{2}$$

Damit ist die allgemeine Lösung der gestellten Aufgabe, das Gesetz gefunden, nach welchem sich die Unsicherheit in der Be-

stimmung der Rechenelemente $x_1, x_2, \ldots x_n$ auf das aus ihnen abgeleitete Resultat f überträgt.

Liegt für jede der Größen $X_1, X_2, \ldots X_n$ nur *eine* Beobachtung $l_1, l_2, \ldots l_n$ vor, so ist $x_1 = l_1$, $x_2 = l_2, \ldots x_n = l_n$ zu setzen, und es bedeutet dann μ_i den mittleren Fehler der Beobachtung l_i.

Ist insbesondere

$$F = X_1 + X_2 + \cdots + X_n, \tag{3}$$

so ist

$$f = x_1 + x_2 + \cdots + x_n,$$

und μ_f, weil die sämtlichen Ableitungen von f den Wert 1 haben, nimmt den Ausdruck an:

$$\mu_f = \sqrt{\mu_1^2 + \mu_2^2 + \cdots + \mu_n^2}, \tag{3*}$$

eine Formel, die auch dann zur Anwendung kommt, wenn mehrere unabhängige Fehlerquellen, die einzeln durch die mittleren Fehler $\mu_1, \mu_2, \ldots \mu_n$ charakterisiert sind, sich zu einem Gesamtfehler vereinigen.

Sind alle Elemente mit gleicher Genauigkeit bestimmt, so daß $\mu_1 = \mu_2 = \cdots = \mu_n = \mu$ ist, so gilt für den letzten Fall die Formel:

$$\mu_f = \mu \sqrt{n}. \tag{3**}$$

Es wächst also der mittlere Fehler der Summe gleich genau bestimmter Summanden wie die Quadratwurzel aus deren Anzahl.

Ist hingegen

$$F = nX, \tag{4}$$

n eine Konstante und x die Bestimmung von X, μ ihr mittlerer Fehler, so hat man

$$f = nx$$

und

$$\mu_f = n\mu. \tag{4*}$$

Der mittlere Fehler eines Vielfachen wächst also wie dieses selbst.

156. Beispiel LIX. a) In Nr. 142b ist die Wahrscheinlichkeit X für das Erscheinen der Nummer 1 aus einer Urne experimentell bestimmt worden, und es ergab sich dafür der Wert

$$x = 0{,}1651$$

mit dem mittleren Fehler

$$\mu = 0{,}0053.$$

Auf Grund dieses Resultates soll die Wahrscheinlichkeit F und ihr mittlerer Fehler bestimmt werden, daß in 6 Ziehungen die Nummer 1 mindestens ein- und höchstens dreimal erscheinen werde.

Es ist

$$F = \binom{6}{3} X^3(1-X)^3 + \binom{6}{2} X^2(1-X)^4 + \binom{6}{1} X(1-X)^5;$$

die vorteilhafteste Bestimmung hierfür ist

$$f = 20x^3(1-x)^3 + 15x^2(1-x)^4 + 6x(1-x)^5$$
$$= x(11x^2 + 3x + 6)(1-x)^3 = 0{,}6530$$

und ihr mittlerer Fehler, da

$$\frac{\partial f}{\partial x} = 60x^2(1-x)^2(1-2x) + 30x(1-x)^3(1-3x)$$
$$+ 6(1-x)^4(1-6x) = 2{,}232$$

ist, kommt gleich

$$\mu_f = \frac{\partial f}{\partial x}\mu = 2{,}232 \cdot 0{,}0053 = 0{,}0119.$$

Das Schlußresultat kann in der Form gegeben werden: Die verlangte Wahrscheinlichkeit hat die mittleren Grenzen

$$0{,}6530 \mp 0{,}0119.$$

b) Die in Nr. 153 besprochene Berliner Grundlinie stellt sich als Summe von zehn Teilstrecken dar; ihre vorteilhafteste Bestimmung ist die Summe der für die Teilstrecken gefundenen Mittelwerte, also

$$2236^m{,}412310;$$

da jede Teilstrecke bei einmaliger Messung mit dem mittleren Fehler $0^{mm}{,}473$ behaftet ist, so ist der mittlere Fehler einer einmaligen Messung der ganzen Grundlinie nach (3**) der vorigen Nummer

$$0^{mm}{,}473\sqrt{10} = 1^{mm}{,}494, \text{ d. i. } \frac{1}{1\,560\,000} \text{ der Länge,}$$

und der mittlere Fehler des arithmetischen Mittels beider Messungen, d. i. des obigen Resultates:

$$\frac{1^{mm}{,}494}{\sqrt{2}} = 1^{mm}{,}057, \text{ d. i. } \frac{1}{2\,210\,000} \text{ der Länge.}$$

c) Im Jahre 1892 wurde bei Bonn eine Grundlinie seitens der Preußischen Landesaufnahme viermal, je zweimal in jeder Richtung, und seitens des Geodätischen Institutes zweimal, je einmal in jeder Richtung, im ganzen also sechsmal gemessen; die Messungen der ersten und zweiten Anstalt erfolgten mit verschiedenen Apparaten. Die Grundlinie wurde dabei nicht in einem Zuge, sondern in 15 durch eingeschaltete Zwischenpunkte hergestellten Abschnitten gemessen.[1])

In der nachstehenden Tabelle sind die Mittelwerte aus den für die einzelnen Strecken gefundenen Längen und ihre mittleren Fehler zusammengestellt; gleichzeitig sind die Quadrate dieser letzteren angeführt.

1) Die Neumessung der Grundlinien bei Strehlen, Berlin und Bonn. Veröffentlichung des Kgl. Preuß. Geodät. Inst., 1897, p. 77—78.

Strecke	Mittelwert der Länge in m	μ in mm	$\mu\mu$
1	234,06008	0,44	0,1936
2	155,97226	0,27	0,0729
3	249,78797	0,33	0,1089
4	155,98596	0,09	0,0081
5	156,10429	0,22	0,0484
6	156,05705	0,36	0,1296
7	156,08547	0,27	0,0729
8	155,97206	0,19	0,0361
9	156,16567	0,32	0,1024
10	156,03199	0,21	0,0441
11	156,13442	0,23	0,0529
12	156,12006	0,25	0,0625
13	155,96778	0,17	0,0289
14	156,06933	0,28	0,0784
15	156,45838	0,35	0,1225
	2512,97277		1,1622 $[\mu\mu]$

Nach Formel (3*) der vorigen Nummer bestimmt sich der mittlere Fehler der Gesamtlänge zu

$$\sqrt{1,1622} = 1^{mm},078, \text{ d. i. } \frac{1}{2\,330\,000} \text{ der Länge.}$$

§ 2. Vermittelnde Beobachtungen.

157. Stellung der Aufgabe. Vorteilhafteste Kombination der Beobachtungen nach dem Prinzip des kleinsten Fehlerrisikos. Eine der unmittelbaren Beobachtung zugängliche Größe V stehe mit den Unbekannten $X, Y, Z, \ldots$, den Elementen[1]) von V, deren Anzahl u sei, in der Beziehung:

$$aX + bY + cZ + \cdots - V = 0.$$

Zu jeder Beobachtung l_ν von V gehöre ein System von Werten

$$a_\nu, \quad b_\nu, \quad c_\nu, \ldots$$

der Koeffizienten, entweder a priori bekannt oder auch durch Beobachtung festgestellt und in diesem Falle als frei von Fehlern angesehen. Ist ε_ν der wahre Fehler von l_ν, so gibt diese Beobachtung Anlaß zu der Gleichung:

$$\varepsilon_\nu = -l_\nu + a_\nu X + b_\nu Y + c_\nu Z + \cdots. \tag{1}$$

Werden n voneinander unabhängige Beobachtungen angestellt, so ergibt sich ein System von n Gleichungen der Form (1).

1) Helmert, Die Ausgleichungsrechnung etc., 2. Aufl., Leipzig 1907, p. 99, nennt die Ausgleichung vermittelnder Beobachtungen auch Elementen-Ausgleichung.

Wäre $n = u$, und setzte man die unbekannten Fehler ε_ν sämtlich gleich Null, so ergäbe sich ein zur Bestimmung von Werten der Unbekannten gerade ausreichendes System. Diese Werte wären aber nicht die wahren, und es gäbe auch kein Mittel, ihre Genauigkeit zu beurteilen.

Ist jedoch $n > u$, dann eröffnet sich die Möglichkeit einer mehrfachen Bestimmung von Werten der Unbekannten; man braucht z. B. nur auf alle möglichen Arten u Gleichungen auszuwählen, ihre linken Seiten zu annullieren und nach den Unbekannten aufzulösen. In der Nichtübereinstimmung der verschiedenen Lösungen äußert sich der Widerspruch, der von der Fehlerhaftigkeit der Beobachtungsresultate $l_1, l_2, \ldots l_n$ herrührt.

Es entsteht daher die Aufgabe, das ganze überzählige System von Gleichungen derart zu kombinieren, daß sich eine einzige Lösung für die Unbekannten ergibt und gleichzeitig gewissen Anforderungen bezüglich der Beschaffenheit dieser Lösung genügt wird.

Bezeichnet man das vorteilhafteste Wertsystem der Unbekannten mit $x, y, z, \ldots$, so führt dasselbe zu analogen Gleichungen, wie das System der wahren Werte, nämlich:

$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z + \cdots \tag{2}$$

mit dem Unterschiede, daß die linken Seiten nicht die wahren, sondern scheinbare Fehler der Beobachtungen vorstellen. Das System (2) muß widerspruchsfrei in dem Sinne sein, daß aus irgend u ausgewählten Gleichungen sich dieselben Werte der $x, y, z, \ldots$ ergeben wie aus jedem andern.

Bei den folgenden Ausführungen beschränken wir uns auf drei Unbekannte, weil die Verallgemeinerung der Resultate auf jede andere Zahl von Unbekannten keine Schwierigkeit darbietet.

Multipliziert man jede Gleichung des Systems (1) mit einem unbestimmten Multiplikator α_ν $(\nu = 1, 2, \ldots n)$ und bildet die Summe, so entsteht die Gleichung:

$$[\alpha\varepsilon] = -[\alpha l] + [a\alpha]X + [b\alpha]Y + [c\alpha]Z;$$

aus ihr ergibt sich, wenn man die Multiplikatoren den Bedingungen

$$[a\alpha] = 1, \quad [b\alpha] = 0, \quad [c\alpha] = 0 \tag{3}$$

unterwirft:

$$X = [\alpha l] + [\alpha\varepsilon]. \tag{4}$$

Die Anwendung eines zweiten Multiplikatorensystems β_ν und eines dritten γ_ν, wenn sie an die Bedingungen

$$\begin{aligned} [a\beta] = 0, \quad [b\beta] = 1, \quad [c\beta] = 0 \\ [a\gamma] = 0, \quad [b\gamma] = 0, \quad [c\gamma] = 1 \end{aligned} \tag{3*}$$

geknüpft werden, führt zu den Gleichungen:

$$\left.\begin{aligned} Y &= [\beta l] + [\beta \varepsilon] \\ Z &= [\gamma l] + [\gamma \varepsilon]. \end{aligned}\right\} \tag{4*}$$

Wird dasselbe Verfahren auf das System (2) angewendet, so kommt man zu den Ansätzen:

$$x = [\alpha l] + [\alpha \lambda] \tag{5}$$

$$\left.\begin{aligned} y &= [\beta l] + [\beta \lambda] \\ z &= [\gamma l] + [\gamma \lambda]. \end{aligned}\right\} \tag{5*}$$

Setzt man zur Abkürzung

$$[\alpha \lambda] = A, \quad [\beta \lambda] = B, \quad [\gamma \lambda] = C, \tag{6}$$

so erhält man durch subtraktive Verbindung von (4), (4*) mit (5), (5*):

$$X - x = -A + [\alpha \varepsilon] \tag{7}$$

$$\left.\begin{aligned} Y - y &= -B + [\beta \varepsilon] \\ Z - z &= -C + [\gamma \varepsilon]. \end{aligned}\right\} \tag{7*}$$

Der Fehler in der Bestimmung x erscheint hiernach als eine lineare Funktion der Beobachtungsfehler. Sind die Beobachtungen *gleich genau*, und befolgen ihre Fehler das symmetrische Gesetz

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2},$$

womit zugleich ausgesprochen ist, daß sie von konstanten Fehleranteilen frei sind, so ist $-A$ der mittlere Wert von $X - x$, daher nach Nr. 148

$$\frac{H}{\sqrt{\pi}} e^{-H^2 (z + A)^2}$$

das Gesetz des Fehlers in der Bestimmung x, wenn

$$H = \frac{h}{\sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}} = \frac{h}{\sqrt{[\alpha \alpha]}} \tag{8}$$

Soll nun das Fehlerrisiko dieser Bestimmung das kleinstmögliche sein, so ist nach Nr. 140 notwendig, daß

$$A = 0 \quad \text{und} \quad [\alpha \alpha] \text{ ein Minimum} \tag{9}$$

sei; dazu kommen die ursprünglich aufgestellten Bedingungen (3).

Die Durchführung dieses relativen Minimums läuft darauf hinaus, das absolute Minimum der mit den unbestimmten Multiplikatoren Q_{11}, Q_{12}, Q_{13} gebildeten Funktion

$$[\alpha \alpha] - 2 Q_{11} ([a \alpha] - 1) - 2 Q_{12} [b \alpha] - 2 Q_{13} [c \alpha]$$

zu bestimmen; dieses aber erfordert, daß die Ableitungen nach den

einzelnen α für sich Null werden. Dadurch ergeben sich die Gleichungen:

$$\alpha_\nu = a_\nu Q_{11} + b_\nu Q_{12} + c_\nu Q_{13} \quad (\nu = 1, 2, \ldots n) \tag{10}$$

Führt man diese Werte in die Bedingungsgleichungen (3) ein, so erhält man zur Bestimmung der Multiplikatoren das Gleichungssystem:

$$\begin{aligned} [aa] Q_{11} + [ab] Q_{12} + [ac] Q_{13} &= 1 \\ [ba] Q_{11} + [bb] Q_{12} + [bc] Q_{13} &= 0 \\ [ca] Q_{11} + [cb] Q_{12} + [cc] Q_{13} &= 0. \end{aligned} \tag{11}$$

Wären Q_{11}, Q_{12}, Q_{13} berechnet, so ergäbe die Einsetzung ihrer Werte in (10) das System der α, das zu dem vorteilhaftesten Werte x führt. Es läßt sich jedoch ein anderer zweckmäßigerer Weg einschlagen.

Vorher möge jedoch noch folgendes bemerkt werden. Multipliziert man (10) nacheinander mit α_ν, β_ν und γ_ν und bildet jedesmal die Summe für alle ν unter Beobachtung der Gleichungen (3) und (3*), so kommt man zu dem Ergebnis, daß

$$Q_{11} = [\alpha\alpha], \quad Q_{12} = [\alpha\beta], \quad Q_{13} = [\alpha\gamma] \tag{12}$$

ist.

Wird die ganze Rechnung in derselben Weise für die beiden andern Unbekannten durchgeführt, so ergeben sich als Bedingungen des kleinsten Fehlerrisikos die Ansätze:

$$\begin{aligned} B = 0 \quad &\text{und} \quad [\beta\beta] \text{ ein Minimum} \\ C = 0 \quad &\text{und} \quad [\gamma\gamma] \text{ ein Minimum} \end{aligned} \tag{9*}$$

daraus für die β und γ die Bestimmungen:

$$\begin{aligned} \beta_\nu &= a_\nu Q_{21} + b_\nu Q_{22} + c_\nu Q_{23} \quad (\nu = 1, 2, \ldots n) \\ \gamma_\nu &= a_\nu Q_{31} + b_\nu Q_{32} + c_\nu Q_{33} \quad (\nu = 1, 2, \ldots n), \end{aligned} \tag{10*}$$

wobei die Multiplikatoren Q_{21}, Q_{22}, Q_{23} und Q_{31}, Q_{32}, Q_{33} zu rechnen sind aus den Gleichungssystemen:

$$\begin{aligned} [aa] Q_{21} + [ab] Q_{22} + [ac] Q_{23} &= 0 \quad & [aa] Q_{31} + [ab] Q_{32} + [ac] Q_{33} &= 0 \\ [ba] Q_{21} + [bb] Q_{22} + [bc] Q_{23} &= 1 \quad & [ba] Q_{31} + [bb] Q_{32} + [bc] Q_{33} &= 0 \\ [ca] Q_{21} + [cb] Q_{22} + [cc] Q_{23} &= 0, \quad & [ca] Q_{31} + [cb] Q_{32} + [cc] Q_{33} &= 1. \end{aligned} \tag{11*}$$

Weiter ergibt sich, daß

$$\begin{aligned} Q_{21} &= [\alpha\beta], \quad Q_{22} = [\beta\beta], \quad Q_{23} = [\beta\gamma] \\ Q_{31} &= [\gamma\alpha], \quad Q_{32} = [\gamma\beta], \quad Q_{33} = [\gamma\gamma] \end{aligned} \tag{12*}$$

Der bloße Anblick der Gleichungen (12) und (12*) zeigt überdies, daß

$$Q_{12} = Q_{21}, \quad Q_{13} = Q_{31}, \quad Q_{23} = Q_{32}.$$

Die Gleichungen

$$A = 0, \quad B = 0, \quad C = 0,$$

d. i. nach (6):

$$[\alpha\lambda] = 0, \quad [\beta\lambda] = 0, \quad [\gamma\lambda] = 0,$$

die nach (9) und (9*) mit zu den Bedingungen des kleinsten Fehlerrisikos gehören, haben drei andere Gleichungen im Gefolge, die sich in folgender Weise ergeben: Multipliziert man die Gleichungen (10) und (10*) mit λ_ν und bildet die Summe für alle ν, so entsteht eben mit Rücksicht darauf, daß $[\alpha\lambda] = 0$, $[\beta\lambda] = 0$ und $[\gamma\lambda] = 0$, das System:

$$\begin{aligned} Q_{11}[a\lambda] + Q_{12}[b\lambda] + Q_{13}[c\lambda] &= 0 \\ Q_{21}[a\lambda] + Q_{22}[b\lambda] + Q_{23}[c\lambda] &= 0 \\ Q_{31}[a\lambda] + Q_{32}[b\lambda] + Q_{33}[c\lambda] &= 0. \end{aligned}$$

Die Determinante aus den Koeffizienten

$$\begin{matrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{matrix} \tag{13}$$

kann nicht Null sein, weil dies im Widerspruche stünde mit dem System

$$\begin{aligned} Q_{11}[aa] + Q_{12}[ab] + Q_{13}[ac] &= 1 \\ Q_{21}[aa] + Q_{22}[ab] + Q_{23}[ac] &= 0 \\ Q_{31}[aa] + Q_{32}[ab] + Q_{33}[ac] &= 0, \end{aligned}$$

das sich aus den ersten Gleichungen der drei Systeme (11) und (11*) zusammensetzt, und in welchen doch $[aa]$, $[ab]$, $[ac]$ von vornherein bestimmte Größen bedeuten. Daher folgt aus (13), daß:

$$[a\lambda] = 0, \quad [b\lambda] = 0, \quad [c\lambda] = 0 \tag{14}$$

sein müsse.

Diese Gleichungen stellen aber die Bedingungen dar, unter welchen die Summe $[\lambda\lambda]$ in bezug auf x, y, z ein Minimum wird; denn nach (2) ist

$$\frac{1}{2}\frac{\partial[\lambda\lambda]}{\partial x} = \left[\lambda\frac{\partial\lambda}{\partial x}\right] = [a\lambda], \text{ usw.}$$

Die vorteilhaftesten, weil mit dem kleinsten Fehlerrisiko verbundenen Werte der Unbekannten sind also diejenigen, welche den Beobachtungen (scheinbare) Fehler von kleinster Quadratsumme zuschreiben.

Damit ist das Prinzip, das der *Methode der kleinsten Quadrate* zugrunde liegt und in Nr. 147 zunächst für direkte Beobachtungen erwiesen wurde, als ein allgemein gültiges Ausgleichungsprinzip be-

gründet. Denn die in Behandlung stehende Aufgabe umfaßt alle Probleme, welche sich der Ausgleichungsrechnung darbieten.

Die Ausführung der Gleichungen (14), nämlich die Einsetzung der Ausdrücke für die λ aus (2), ergibt das Gleichungssystem:

$$\begin{aligned} [aa]x + [ab]y + [ac]z &= [al] \\ [ba]x + [bb]y + [bc]z &= [bl] \\ [ca]x + [cb]y + [cc]z &= [cl], \end{aligned} \tag{15}$$

durch das die vorteilhaftesten Werte der Unbekannten eindeutig bestimmt sind, sofern die Determinante der Koeffizienten nicht verschwindet. Es heißt das System der *Normalgleichungen*.

158. Gewichte der Unbekannten. Weil bei der vorteilhaftesten Kombination die Größen A, B, C der vorigen Nummer Null sind, so schreiben sich die Gleichungen (7) und (7*):

$$\begin{aligned} X - x &= [\alpha\,\varepsilon] \\ Y - y &= [\beta\,\varepsilon] \\ Z - z &= [\gamma\,\varepsilon]. \end{aligned}$$

Ist daher h das Präzisionsmaß einer Beobachtung, so ist nach Nr. 138

$$H_x = \frac{h}{\sqrt{[\alpha\alpha]}}$$

das Präzisionsmaß in der Bestimmung von x; ebenso sind

$$H_y = \frac{h}{\sqrt{[\beta\beta]}}, \qquad H_z = \frac{h}{\sqrt{[\gamma\gamma]}}$$

die Präzisionsmaße von y und z.

Führt man an Stelle der Präzisionsmaße die mittleren Fehler ein, wobei die Formeln

$$h = \frac{1}{\mu\sqrt{2}}, \qquad H_x = \frac{1}{\mu_x\sqrt{2}} \quad \text{usw.}$$

zu benützen sind, so ergeben sich für die mittleren Fehler der Unbekannten die Formeln:

$$\mu_x = \mu\sqrt{[\alpha\alpha]}, \qquad \mu_y = \sqrt{[\beta\beta]}, \qquad \mu_z = \mu\sqrt{[\gamma\gamma]}, \tag{16}$$

ausgedrückt durch den mittleren Fehler μ einer Beobachtung. Wird diese als Gewichtseinheit genommen, so kommen den Bestimmungen der Unbekannten folgende Gewichte zu (s. Nr. 149):

$$p_x = \frac{\mu^2}{\mu_x^2} = \frac{1}{[\alpha\alpha]}, \qquad p_y = \frac{\mu^2}{\mu_y^2} = \frac{1}{[\beta\beta]}, \qquad p_z = \frac{\mu^2}{\mu^2} = \frac{1}{[\gamma\gamma]} \tag{17}$$

Da nun bei der vorteilhaftesten Kombination

$$[\alpha\alpha], \quad [\beta\beta], \quad [\gamma\gamma] \tag{18}$$

Minima sind, so haben die nach der Methode der kleinsten Quadrate bestimmten Werte der Unbekannten zugleich die *größten Gewichte.*

Nach den Gleichungen (12) und (12*) sind die Summen (18) gleichbedeutend mit den Größen

$$Q_{11}, \quad Q_{22}, \quad Q_{33}, \tag{18*}$$

zu deren Bestimmung die Gleichungssysteme (11) und (11*) dienen. Diese Gleichungssysteme werden deshalb als die *Gewichtsgleichungen* bezeichnet, und zwar bilden (11) die Gewichtsgleichungen für die Unbekannte x, weil sie Q_{11} enthalten, (11*) die Gewichtsgleichungen für y und z.

Man bemerkt, daß die Gewichtsgleichungen mit den Normalgleichungen (15) übereinstimmend gebaut sind und nur in den rechten Seiten sich von ihnen unterscheiden.

159. Bestimmung des mittleren Fehlers einer Beobachtung aus den scheinbaren Fehlern. Sobald man diese Aufgabe durchgeführt und auch die Gewichte der Unbekannten berechnet hat, sind nach den Ausführungen der vorigen Nummer auch die mittleren Fehler der letzteren bestimmbar.

Zur Lösung der in der Überschrift bezeichneten Aufgabe benützen wir das von Gauß[1]) angegebene Verfahren.

Geht man von den Gleichungen (1), (2) aus, die da lauten:

$$\varepsilon_\nu = -l_\nu + a_\nu X + b_\nu Y + c_\nu Z$$
$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z,$$

so ergibt sich durch deren Subtraktion:

$$\lambda_\nu = \varepsilon_\nu - a_\nu(X - x) - b_\nu(Y - y) - c_\nu(Z - z);$$

den Gleichungen (7) und (7*) zufolge ist aber bei der vorteilhaftesten Kombination wegen $A = 0,\ B = 0,\ C = 0$:

$$X - x = [\alpha\varepsilon], \quad Y - y = [\beta\varepsilon], \quad Z - z = [\gamma\varepsilon];$$

hiernach ist

$$\lambda_\nu = \varepsilon_\nu - a_\nu[\alpha\varepsilon] - b_\nu[\beta\varepsilon] - c_\nu[\gamma\varepsilon],$$

also λ_ν als homogene lineare Funktion der wahren Fehler dargestellt. Multipliziert man diese Gleichung mit ε_ν und bildet die Summe für alle ν von 1 bis n, so entsteht:

$$[\lambda\varepsilon] = [\varepsilon\varepsilon] - [a\varepsilon][\alpha\varepsilon] - [b\varepsilon][\beta\varepsilon] - [c\varepsilon][\gamma\varepsilon];$$

---

1) Theoria combinat. Art. 38. — Andere Lösungen siehe in des Verf.s „Theorie der Beobachtungsfehler" und bei P. Pizzetti, l. c., p. 262ff.

multipliziert man hingegen mit λ_ν und bildet wieder die Summe, so kommt man mit Rücksicht darauf, daß nach (14)

$$[a\lambda] = 0, \quad [b\lambda] = 0, \quad [c\lambda] = 0$$

ist, zu der Gleichung

$$[\lambda\lambda] = [\lambda\varepsilon],$$

welche mit der letzten verbunden eine Darstellung von $[\lambda\lambda]$ als homogene quadratische Funktion der wahren Fehler ergibt, nämlich:

$$[\lambda\lambda] = [\varepsilon\varepsilon] - [a\varepsilon][\alpha\varepsilon] - [b\varepsilon][\beta\varepsilon] - [c\varepsilon][\gamma\varepsilon]. \tag{19}$$

Nach Auflösung der Summen ist:

$$\begin{aligned}
[\lambda\lambda] = [\varepsilon\varepsilon] & - \varepsilon_1^2(a_1\alpha_1 + b_1\beta_1 + c_1\gamma_1) \\
& - \varepsilon_2^2(a_2\alpha_2 + b_2\beta_2 + c_2\gamma_2) \\
& \ldots\ldots\ldots\ldots \\
& - \varepsilon_n^2(a_n\alpha_n + b_n\beta_n + c_n\gamma_n) \\
& - \varepsilon_1\varepsilon_2(a_1\alpha_2 + a_2\alpha_1 + b_1\beta_2 + b_2\beta_1 + c_1\gamma_2 + c_2\gamma_1) \\
& - \varepsilon_1\varepsilon_3(a_1\alpha_3 + a_3\alpha_1 + b_1\beta_3 + b_3\beta_1 + c_1\gamma_3 + c_3\gamma_1) \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& - \varepsilon_{n-1}\varepsilon_n(a_{n-1}\alpha_n + a_n\alpha_{n-1} + b_{n-1} + \beta_n b_n\beta_{n-1} + c_{n-1}\gamma_n + c_n\gamma_{n-1}).
\end{aligned}$$

Man ersetze nun auf der rechten Seite jedes Glied durch seinen Mittelwert und betrachte den so erhaltenen Wert des rechtsseitigen Ausdruckes als gleichwertig mit $[\lambda\lambda]$. Dabei ist zu beachten, daß der Mittelwert eines jeden ε_ν^2 gleich ist μ^2, und der Mittelwert eines jeden $\varepsilon_i\varepsilon_k$ gleich Null, weil vorausgesetzt wurde, daß die Fehler um die Null symmetrisch angeordnet oder frei von einem konstanten Anteil sind. Demnach ergibt dieser Vorgang den Ansatz:

$$\begin{aligned}
[\lambda\lambda] = n\mu^2 & - \mu^2(a_1\alpha_1 + b_1\beta_1 + c_1\gamma_1) \\
& - \mu_2(a_2\alpha_2 + b_2\alpha_2 + c_2\gamma_2) \\
& \ldots\ldots\ldots \\
& - \mu^2(a_n\alpha_n + b_n\beta_n + c_n\gamma_n) \\
= n\mu^2 & - \mu^2 \{ [a\alpha] + [b\beta] + [c\gamma] \};
\end{aligned}$$

nach den grundlegenden Bedingungen (3) und (3*) ist aber

$$[a\alpha] = 1, \quad [b\beta] = 1, \quad [c\gamma] = 1,$$

daher

$$[\lambda\lambda] = (n-3)\mu^2$$

und daraus

$$\mu = \sqrt{\frac{[\lambda\lambda]}{n-3}}.$$

Diese Formel kam unter der Voraussetzung zustande, daß drei Unbekannte zu bestimmen seien; allgemein ist für u Unbekannte:

$$\mu = \sqrt{\frac{[\lambda\lambda]}{n-u}}. \tag{20}$$

160. Vermittelnde Beobachtungen ungleicher Genauigkeit. Sind die Beobachtungen $l_1, l_2, \cdots l_n$ von ungleicher Genauigkeit, und ist diese durch die Präzisionsmaße $h_1, h_2, \cdots h_n$ bezeichnet, so befolgt der Fehler ε_ν der Beobachtung l_ν das Gesetz

$$\frac{h_\nu}{\sqrt{\pi}} e^{-h_\nu^2 \varepsilon_\nu^2}; \tag{1}$$

es wird also das Produkt $h_\nu \varepsilon_\nu = \varepsilon_\nu'$ das Gesetz

$$\frac{1}{\sqrt{\pi}} e^{-\varepsilon_\nu'^2} \tag{1*}$$

befolgen (s. Nr. 138). Denkt man sich also jede der Gleichungen

$$\varepsilon_\nu = -l_\nu + a_\nu X + b_\nu Y + c_\nu Z \tag{2}$$

wie auch die ihr zugeordnete

$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z \tag{3}$$

mit h_ν multipliziert, so entstehen neue Gleichungssysteme:

$$\varepsilon_\nu' = -l_\nu' + a_\nu' X + b_\nu' Y + c_\nu' Z \tag{2*}$$
$$(\nu = 1, 2, \ldots n),$$
$$\lambda_\nu' = -l_\nu' + a_\nu' x + b_\nu' y + c_\nu' z \tag{3*}$$

die sich auf ein einheitliches Gesetz (1*) beziehen und daher ebenso zu behandeln sind wie die von Nr. 157.

Es ist jedoch üblich, das Genauigkeitsverhältnis der Beobachtungen entweder durch ihre mittleren Fehler oder durch die Gewichte zum Ausdruck zu bringen. Nach Nr. 149 sind aber die Präzisionsmaße den mittleren Fehlern umgekehrt, den Quadratwurzeln aus den Gewichten direkt proportional. Statt also die Gleichungen (2), (3), um sie, wie man sich ausdrückt, *auf gleiches Gewicht zu reduzieren*, mit den Präzisionsmaßen zu multiplizieren, kann man sie durch die mittleren Fehler der bezüglichen Beobachtungen dividieren oder mit den Quadratwurzeln aus den korrespondierenden Gewichten multiplizieren.

Wir wollen den letzteren Vorgang als den üblichen verfolgen und bezeichnen die Gewichte von $l_1, l_2, \cdots l_n$ mit $p_1, p_2, \cdots p_n$. Befolgt der Fehler ε_ν das Gesetz (1), so unterliegt $\sqrt{p_\nu}\varepsilon_\nu = \varepsilon_\nu'$ nach Nr. 138 dem Gesetz

$$\frac{\frac{h_\nu}{\sqrt{p_\nu}}}{\sqrt{\pi}} e^{-\frac{h_\nu^2}{p_\nu}\varepsilon_\nu'^2},$$

und da sich alle Gewichte auf dieselbe Gewichtseinheit beziehen müssen, so ist

$$\frac{h_\nu}{\sqrt{p_\nu}} = h$$

unabhängig von ν und bezeichnet das Präzisionsmaß der *Gewichtseinheit.* Es unterliegen also bei diesem Vorgange die Gleichungen (2*) dem einheitlichen Gesetze

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon_\nu'^2},$$

dem die Gewichtseinheit unterworfen ist.

Die nun geltenden *Normalgleichungen*:

$$\begin{aligned} [a'a']x + [a'b']y + [a'c']z &= [a'l'] \\ [b'a']x + [b'b']y + [b'c']z &= [b'l'] \\ [c'a']x + [c'b']y + [c'c']z &= [c'l'] \end{aligned} \tag{4}$$

lauten in den ursprünglichen Größen:

$$\begin{aligned} [paa]x + [pab]y + [pac]z &= [pal] \\ [pba]x + [pbb]y + [pbc]z &= [pbl] \\ [pca]x + [pcb]y + [pcc]z &= [pcl]; \end{aligned} \tag{4*}$$

ebenso gestalten sich die *Gewichtsgleichungen* für x in den alten Größen wie folgt:

$$\begin{aligned} [paa]Q_{11} + [pab]Q_{12} + [pac]Q_{13} &= 1 \\ [pba]Q_{11} + [pbb]Q_{12} + [pbc]Q_{13} &= 0 \\ [pca]Q_{11} + [pcb]Q_{12} + [pcc]Q_{13} &= 0 \end{aligned} \tag{5*}$$

und ähnlich die für y und z.

Die Gleichungen (14) in Nr. 157:

$$[a\lambda] = 0, \quad [b\lambda] = 0, \quad [c\lambda] = 0,$$

aus denen die Normalgleichungen hervorgegangen sind, lauten nunmehr:

$$[pa\lambda] = 0, \quad [pb\lambda] = 0, \quad [pc\lambda] = 0 \tag{6*}$$

und stellen die Bedingungen vor, unter welchen die Summe $[p\lambda\lambda]$ in bezug auf x, y, z ein Minimum erlangt. Daraus ergibt sich die Erweiterung des in Nr. 157 formulierten Prinzips auf den Fall ungleich genauer Beobachtungen. *Es sind* nämlich dann *jene Werte der Unbekannten die vorteilhaftesten, für welche die Summe der mit den Gewichten multiplizierten Quadrate der (scheinbaren) Fehler am kleinsten wird.*

Für den *mittleren Fehler der Gewichtseinheit* ergibt sich aus den auf gleiches Gewicht (= 1) reduzierten λ' der Ausdruck:

$$\mu = \sqrt{\frac{[\lambda'\lambda']}{n-u}}, \tag{7}$$

wenn u die Anzahl der Unbekannten ist; in den ursprünglichen λ ausgedrückt lautet diese Formel:

$$\mu = \sqrt{\frac{[p\lambda\lambda]}{n-u}}. \tag{7*}$$

Aus μ ergeben sich mittels der Gewichte der Unbekannten deren mittlere Fehler nach den Formeln:

$$\mu_x = \frac{\mu}{\sqrt{p_x}} = \mu\sqrt{Q_{11}}, \quad \mu_y = \frac{\mu}{\sqrt{p_y}} = \mu\sqrt{Q_{22}} \text{ usw.}$$

161. Auflösung der Normalgleichungen und der Gewichtsgleichungen. Hierbei ist es nicht notwendig, den Fall ungleich genauer Beobachtungen von dem gleich genauer zu unterscheiden, weil Normal- und Gewichtsgleichungen in beiden Fällen dieselbe Form haben.

Wenn es sich nur um wenige Unbekannte handelt und die Koeffizienten der Gleichungen einfache Zahlen sind, dann wird man sich eines der verschiedenen Kombinationsverfahren der Algebra bedienen, um möglichst rasch zum Ziele zu kommen. Sobald aber die Anzahl der Unbekannten einigermaßen erheblich ist und die Koeffizienten vielstellige Zahlen sind, empfiehlt sich jener systematische Vorgang, den Gauß[1]) angegeben hat.

Aus der ersten Gleichung des Systems

$$\begin{aligned} [aa]x + [ab]y + [ac]z &= [al] \\ [ab]x + [bb]y + [bc]z &= [bl] \\ [ac]x + [bc]y + [cc]z &= [cl] \end{aligned} \tag{1}$$

folgt:

$$x + \frac{[ab]}{[aa]}y + \frac{[ac]}{[aa]}z = \frac{[al]}{[aa]}. \tag{2}$$

Subtrahiert man diese Gleichung, nachdem man sie mit $[ab]$, beziehungsweise $[ac]$ multipliziert hat, von der zweiten und dritten, so entsteht das Gleichungspaar:

$$\left\{[bb] - \frac{[ab]}{[aa]}[ab]\right\}y + \left\{[bc] - \frac{[ab]}{[aa]}[ac]\right\}z = [bl] - \frac{[ab]}{[aa]}[al]$$

$$\left\{[bc] - \frac{[ab]}{[aa]}[ac]\right\}y + \left\{[cc] - \frac{[ac]}{[aa]}[ac]\right\}z = [cl] - \frac{[ac]}{[aa]}[al],$$

---

1) Disquisitio de elementis ellipticis Palladis. Comment. Gott. 1810, und Theoriae combin. suppl. 1826, Art. 13.

das analog gebaut ist wie das System (1); denn die zur Hauptdiagonale symmetrisch liegenden Koeffizienten sind wieder gleich und die der Hauptdiagonale selbst Quadratsummen wie dort, daher positiv; so ist beispielsweise

$$[bb]-\frac{[ab]}{[aa]}[ab]=\frac{[aa][bb]-[ab]^2}{[aa]}=\left(\frac{a_1 b_2-a_2 b_1}{\sqrt{[aa]}}\right)^2+\left(\frac{a_1 b_3-a_3 b_1}{\sqrt{[aa]}}\right)^2+\cdots.$$

Ferner zeigen die Koeffizienten und die absoluten Glieder einen einheitlichen Bau, der sich durch das Schema

$$[ik]-\frac{[ai]}{[aa]}[ak]$$

darstellen läßt; führt man dafür das Symbol

$$[ik\cdot 1]$$

ein, wonach also

$$[bb]-\frac{[ab]}{[aa]}[ab]=[bb\cdot 1],\quad [bc]-\frac{[ab]}{[aa]}[ac]=[bc\cdot 1],\quad\cdots \tag{3}$$

ist, so schreibt sich jenes Gleichungspaar wie folgt:

$$\begin{aligned}[bb\cdot 1]y+[bc\cdot 1]z&=[bl\cdot 1]\\ [bc\cdot 1]y+[cc\cdot 1]z&=[cl\cdot 1].\end{aligned} \tag{4}$$

Aus der ersten dieser Gleichungen folgt:

$$y+\frac{[bc\cdot 1]}{[bb\cdot 1]}z=\frac{[bl\cdot 1]}{[bb\cdot 1]}; \tag{5}$$

multipliziert man dies mit $[bc\cdot 1]$ und subtrahiert von der zweiten Gleichung, so entsteht:

$$\left\{[cc\cdot 1]-\frac{[bc\cdot 1]}{[bb\cdot 1]}[bc\cdot 1]\right\}z=[cl\cdot 1]-\frac{[bc\cdot 1]}{[bb\cdot 1]}[bl\cdot 1];$$

Koeffizient und rechte Seite sind hier nach dem Schema

$$[ik\cdot 1]-\frac{[bi\cdot 1]}{[bb\cdot 1]}[bk\cdot 1]$$

gebildet; gebraucht man hierfür die Abkürzung

$$[ik\cdot 2],$$

wonach also

$$[cc\cdot 1]-\frac{[bc\cdot 1]}{[bb\cdot 1]}[bc\cdot 1]=[cc\cdot 2],\quad\cdots \tag{6}$$

so schreibt sich die obige Gleichung:

$$[cc\cdot 2]z=[cl\cdot 2]. \tag{7}$$

Aus ihr folgt dann:

$$z=\frac{[cl\cdot 2]}{[cc\cdot 2]}. \tag{8}$$

Die Gleichungen (2), (5), (8) sind geeignet, das System (1) der Normalgleichungen zu ersetzen; ihre Vereinigung:

$$
\begin{aligned}
x + \frac{[ab]}{[aa]} y + \frac{[ac]}{[aa]} z &= \frac{[al]}{[aa]} \\
y + \frac{[bc \cdot 1]}{[bb \cdot 1]} z &= \frac{[bl \cdot 1]}{[bb \cdot 1]} \\
z &= \frac{[cl \cdot 2]}{[cc \cdot 2]}
\end{aligned}
\qquad . \qquad (9)
$$

bildet das *System der reduzierten Normalgleichungen.* Dasselbe gestattet die Berechnung der Unbekannten durch sukzessive Substitution von der letzten zur ersten. Hat man die darin auftretenden Brüche, welche in ihrer Reihenfolge kurz mit

$$
\begin{matrix}
\mathfrak{b}_1 & \mathfrak{c}_1 & \mathfrak{l}_1 \\
 & \mathfrak{c}_2 & \mathfrak{l}_2 \\
 & & \mathfrak{l}_3
\end{matrix}
$$

bezeichnet werden mögen, berechnet, so lassen sich mit Hilfe derselben x, y, z explizite darstellen; es ist nämlich

$$
\begin{aligned}
x &= \begin{vmatrix} \mathfrak{l}_1 & \mathfrak{b}_1 & \mathfrak{c}_1 \\ \mathfrak{l}_2 & 1 & \mathfrak{c}_2 \\ \mathfrak{l}_3 & 0 & 1 \end{vmatrix} = \mathfrak{l}_1 - \mathfrak{b}_1 \mathfrak{l}_2 + (\mathfrak{b}_1 \mathfrak{c}_2 - \mathfrak{c}_1)\mathfrak{l}_3 \\
y &= \begin{vmatrix} \mathfrak{l}_2 & \mathfrak{c}_2 \\ \mathfrak{l}_3 & 1 \end{vmatrix} = \mathfrak{l}_2 - \mathfrak{c}_2 \mathfrak{l}_3 \\
z &= \mathfrak{l}_3 .
\end{aligned}
\qquad (10)
$$

Würde man dasselbe Eliminationsverfahren, das hier auf die Normalgleichungen ausgeübt worden ist, auf die Gewichtsgleichungen von z, d. i. auf

$$
\begin{aligned}
[aa] Q_{31} + [ab] Q_{32} + [ac] Q_{33} &= 0 \\
[ab] Q_{31} + [bb] Q_{32} + [bc] Q_{33} &= 0 \\
[ac] Q_{31} + [bc] Q_{32} + [cc] Q_{33} &= 1
\end{aligned}
\qquad (11)
$$

zur Anwendung bringen, so ergäben sich in (9) links dieselben Koeffizienten, rechts dagegen andere Werte; es entsteht nämlich (11) aus (1), wenn man außer der Änderung der Zeichen für die Unbekannten

$$[al] \text{ durch } 0, \quad [bl] \text{ durch } 0, \quad [cl] \text{ durch } 1$$

ersetzt; infolgedessen wird weiter

$$
\begin{aligned}
[bl \cdot 1] &= [bl] - \frac{[ab]}{[aa]} [al] \text{ gleich } 0 \\
[cl \cdot 1] &= [cl] - \frac{[ac]}{[aa]} [al] \text{ gleich } 1 \\
[cl \cdot 2] &= [cl \cdot 1] - \frac{[bc \cdot 1]}{[bb \cdot 1]} [bl \cdot 1] \text{ gleich } 1 ,
\end{aligned}
$$

und das zu (11) gehörige reduzierte System lautet:

$$
\begin{aligned}
Q_{31} + \frac{[ab]}{[aa]} Q_{32} + \frac{[ac]}{[aa]} Q_{33} &= 0 \\
Q_{32} + \frac{[bc\cdot 1]}{[bb\cdot 1]} Q_{33} &= 0 \\
Q_{33} &= \frac{1}{[cc\cdot 2]}\,.
\end{aligned}
\tag{12}
$$

Am bemerkenswertesten ist die dritte dieser Gleichungen, da sie aussagt, daß bei dem eingeschlagenen Eliminationsverfahren *der Nenner der letzten Unbekannten zugleich ihr Gewicht* angibt.

Auf diese Weise ist aber eben nur das Gewicht der letzten Unbekannten einfach gefunden. Um das der ersten zu erhalten, kann man das System der Normalgleichungen derart *umstellen*, daß die Gleichungen *und* die Unbekannten in die umgekehrte Ordnung kommen; führt man bei dieser Anordnung das Eliminationsverfahren wie früher durch, so gibt der Nenner von x zugleich dessen Gewicht. Zur Bestimmung des Gewichtes von y wäre nur die Umstellung des abgekürzten Systemes (4) erforderlich. Wählt man diesen Vorgang, so gewährt die mehrfache Bestimmung der Unbekannten zugleich eine Kontrolle der Rechnung.

Bei einer größeren Anzahl von Unbekannten wird jedoch eine solche Wiederholung des Eliminationsverfahrens beschwerlich; es empfiehlt sich dann, jedes System von Gewichtsgleichungen ebenso zu reduzieren wie das Normalgleichungssystem. Dabei kann von dem Umstande, daß $Q_{ik} = Q_{ki}$, vorteilhaft Gebrauch gemacht werden. Hat man das System (12) vollständig aufgelöst, also Q_{31}, Q_{32}, Q_{33} berechnet, so ist für das System der zu y gehörigen Gewichtsgleichungen eine Unbekannte, nämlich Q_{23}, schon bestimmt; man kann sich daher auf die ersten zwei Gleichungen:

$$
\begin{aligned}
[aa]\,Q_{21} + [ab]\,Q_{22} + [ac]\,Q_{23} &= 0 \\
[ba]\,Q_{21} + [bb]\,Q_{22} + [bc]\,Q_{23} &= 1
\end{aligned}
$$

beschränken, sie auf die reduzierte Form

$$
\begin{aligned}
Q_{21} + \frac{[ab]}{[aa]} Q_{22} + \frac{[ac]}{[aa]} Q_{23} &= 0 \\
Q_{22} + \frac{[bc\cdot 1]}{[bb\cdot 1]} Q_{23} &= \frac{1}{[bb\cdot 1]}
\end{aligned}
\tag{13}
$$

bringen und nach Q_{21}, Q_{22} auflösen. Dann aber sind für das zu x gehörige Gewichtsgleichungssystem bereits zwei Unbekannte, Q_{12} und Q_{13}, bestimmt, und man kann sich auf die eine Gleichung

$$
[aa]\,Q_{11} + [ab]\,Q_{12} + [ac]\,Q_{13} = 1
$$

beschränken, die man zum Zwecke der Berechnung von Q_{11} auf die reduzierte Form

$$Q_{11} + \frac{[ab]}{[aa]} Q_{12} + \frac{[ac]}{[aa]} Q_{13} = \frac{1}{[aa]} \tag{14}$$

bringen wird. Der Umstand, daß man es in den Gleichungen (12), (13), (14) auf der linken Seite immer mit denselben Koeffizienten zu tun hat, bedeutet eine wesentliche Vereinfachung der Rechnungen.[1])

162. Nichtlineare Relationen zwischen der beobachteten und den unbekannten Größen. Die Theorie der Ausgleichung vermittelnder Beobachtungen bedarf noch einer wichtigen Ergänzung. In Nr. 157 ist nämlich angenommen worden, daß zwischen der zu beobachtenden Größe V und den Unbekannten $X, Y, Z, \cdots$ eine lineare Beziehung bestehe. Ist dem nicht so, ist der Zusammenhang ein anderer, so ist eine vorbereitende Rechnung erforderlich, bevor an die Ausgleichungsrechnung selbst geschritten werden kann.

Es bestehe allgemein zwischen V und den (drei) Elementen die Beziehung:

$$0 = -V + F(X, Y, Z). \tag{1}$$

Die Beobachtung l_ν von V, mit dem Fehler ε_ν, führt zu der Gleichung:

$$\varepsilon_\nu = -l_\nu + F_\nu(X, Y, Z); \tag{2}$$

der Zeiger ν bei F deutet auf Parameter hin, welche sich von Beobachtung zu Beobachtung ändern.

Wählt man in zweckmäßiger Weise so viel Beobachtungen aus, als es Unbekannte gibt, und betrachtet sie als fehlerfrei, so ergeben sich Gleichungen in gerade zureichender Zahl zur Bestimmung von *Näherungswerten* X_0, Y_0, Z_0 der Unbekannten. Von diesen werde angenommen, daß die *Korrektionen* ξ, η, ζ, durch welche sie auf die wahren Werte X, Y, Z ergänzt werden, so klein seien, daß man Glieder mit Potenzen und Produkten derselben außer acht lassen kann. Unter solchen Verhältnissen darf

$$\begin{aligned} F_\nu(X, Y, Z) &= F_\nu(X_0 + \xi, Y_0 + \eta, Z_0 + \zeta) \\ &= F_\nu(X_0, Y_0, Z_0) + \frac{\partial F_\nu}{\partial X_0}\xi + \frac{\partial F_\nu}{\partial Y_0}\eta + \frac{\partial F_\nu}{\partial Z_0}\zeta \end{aligned}$$

gesetzt werden, und es geht die Gleichung (2) über in:

$$\varepsilon_\nu = -l_\nu + F_\nu(X_0, Y_0, Z_0) + \frac{\partial F_\nu}{\partial X_0}\xi + \frac{\partial F_\nu}{\partial Y_0}\eta + \frac{\partial F_\nu}{\partial Z_0}\zeta. \tag{3}$$

---

1) Über Methoden zur indirekten Auflösung der Normal- und Gewichtsgleichungen, durch die eine fortgesetzte Annäherung an die Werte der Unbekannten angestrebt wird, sehe man Helmert, Die Ausgleichungsrechnung usw., 2. Aufl., Leipzig 1907, p. 175—180.

Wenn nun

$$\begin{aligned} l_\nu - F_\nu(X_0, Y_0, Z_0) &= l'_\nu \\ \frac{\partial F_\nu}{\partial X_0} &= a_\nu \\ \frac{\partial F_\nu}{\partial Y_0} &= b_\nu \\ \frac{\partial F_\nu}{\partial Z_0} &= c_\nu \end{aligned} \tag{4}$$

gesetzt wird, so nimmt (3) die in Nr. 157 vorausgesetzte lineare Form an:

$$\varepsilon_\nu = -l'_\nu + a_\nu \xi + b_\nu \eta + c_\nu \zeta. \tag{5}$$

Jeder solchen Gleichung entspricht eine andere:

$$\lambda_\nu = -l'_\nu + a_\nu x + b_\nu y + c_\nu z, \tag{6}$$

in der x, y, z die vorteilhaftesten Werte der Korrektionen bedeuten, so daß

$$X_0 + x, \quad Y_0 + y, \quad Z_0 + z \tag{7}$$

die vorteilhaftesten Werte der Unbekannten selbst sind.

Aber erst der Erfolg der Rechnung kann darüber belehren, ob die Voraussetzung zulässig war, auf welcher der Ansatz (3) beruhte. Sollten sich für x, y, z so große Werte ergeben, daß man gegen die Außerachtlassung der Potenzen und Produkte Bedenken tragen muß, dann wäre die ganze Rechnung mit den Werten (7) als *neuen Näherungswerten* zu wiederholen.

Die Einführung von Näherungswerten für die Unbekannten empfiehlt sich zumeist auch dann, wenn die Relation zwischen V und $X, Y, Z, \cdots$ von Natur aus linear ist; sie hat dann aber einen andern Zweck als vorhin, nämlich den der *Vereinfachung der numerischen Rechnung.* Sind nämlich die Beobachtungsergebnisse l_ν vielziffrige Zahlen, so kann durch passende Näherungswerte bewirkt werden, daß die nach der Vorschrift (4) reduzierten Beobachtungen:

$$l'_\nu = l_\nu - F_\nu(X_0, Y_0, Z_0) = l_\nu - (a_\nu X_0 + b_\nu Y_0 + c_\nu Z_0)$$

bequemere Zahlen werden.

Mitunter kann die Umwandlung der nichtlinearen Relation (1) in eine lineare auch auf einem andern Wege, mit Umgehung von Näherungswerten, bewerkstelligt werden, wenn sich eine analytische Operation f angeben läßt, durch welche die Funktion F in eine lineare umgewandelt wird[1]), so daß

1) Als Beispiel diene: $V = Ce^{aX+bY}$; durch die Operation des Logarithmierens im natürlichen System geht daraus die in bezug auf X, Y lineare Beziehung: $l \cdot V = l \cdot C + aX + bY$ hervor.

$$f(V) = f\{F(X, Y, Z)\} = K + aX + bY + cZ \tag{8}$$

ist. Nun muß aber beachtet werden, daß die Ausgleichung nicht an den Beobachtungen von V selbst, sondern an einer Funktion $f(V)$ der Beobachtungswerte vorgenommen wird, und dies hat auf die Gewichte der Gleichungen (8) Einfluß.

Ist l_ν eine Beobachtung von V, ε_ν ihr Fehler; gehören ferner zu dieser Beobachtung die Werte a_ν, b_ν, c_ν der Parameter, so gibt sie zu der Gleichung:

$$f(l_\nu + \varepsilon_\nu) = K_\nu + a_\nu X + b_\nu Y + c_\nu Z \tag{9}$$

Anlaß; für ein sehr kleines ε_ν kann

$$f(l_\nu + \varepsilon_\nu) = f(l_\nu) + \frac{\partial f}{\partial l_\nu}\varepsilon_\nu$$

gesetzt werden. Hat nun l_ν die Präzision h_ν, das Gewicht p_ν, so hat nach Nr. 138 wegen

$$f(l_\nu + \varepsilon_\nu) - f(l_\nu) = \frac{\partial f}{\partial l_\nu}\varepsilon_\nu$$

$f(l_\nu)$ die Präzision

$$h_\nu' = \frac{h_\nu}{\frac{\partial f}{\partial l_\nu}}$$

und ein Gewicht p_ν', das sich aus der Beziehung

$$p_\nu' : p_\nu = \frac{h_\nu^2}{\left(\frac{\partial f}{\partial l_\nu}\right)^2} : h_\nu^2$$

ergibt; somit ist

$$p_\nu' = \frac{p_\nu}{\left(\frac{\partial f}{\partial l_\nu}\right)^2}. \tag{10}$$

Waren die Beobachtungen l_ν von gleicher Genauigkeit und vom Gewicht 1, so bekommt die Gleichung

$$\varepsilon_\nu' = -l_\nu' + a_\nu X + b_\nu Y + c_\nu Z, \tag{11}$$

welche aus (9) hervorgeht, wenn man

$$\frac{df}{dl_\nu}\varepsilon_\nu = \varepsilon_\nu', \quad -f(l_\nu) + K_\nu = l_\nu'$$

setzt, das Gewicht

$$p_\nu' = \frac{1}{\left(\frac{\partial f}{\partial l_\nu}\right)^2}. \tag{10*}$$

Auch diese Methode[1]), die nur in ganz besonderen Fällen an-

1) Auf ihre korrekte Durchführung hat Th. Wittstein hingewiesen Zeitschr. f. Math. u. Phys. XXVII (1882), p. 315, und „Das mathematische Gesetz der Sterblichkeit", 2. Aufl., 1883.

wendbar ist, hängt von einer Voraussetzung, nämlich von solcher Kleinheit des Beobachtungsfehlers ab, daß dessen Potenzen vernachlässigt werden können.

163. Zusammenstellung der Resultate. Das *erste* Geschäft bei der Ausgleichung vermittelnder Beobachtungen ist die *Aufstellung der Fehlergleichungen*, die schließlich immer die Form:

$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z \tag{1}$$

annehmen. Ihre Anzahl stimmt mit der Anzahl der Beobachtungen überein. Handelt es sich um lineare Beziehungen, so bedarf dies keiner besonderen Rechnung; nur bei Einführung von Näherungswerten X_0, Y_0, Z_0 zum Zwecke der Reduktion der l hat man für jede Gleichung den Wert von

$$a_\nu X_0 + b_\nu Y_0 + c_\nu Z_0$$

zu berechnen, von l_ν zu subtrahieren und mit den Differenzen l'_ν wieder Gleichungen von der Form (1) zu bilden, in welchen x, y, z nunmehr Korrektionen an den Näherungswerten bedeuten. Bei nichtlinearen Relationen erfordert die Aufstellung der Fehlergleichungen besondere Vorarbeiten, die in Nr. 162 erörtert sind.

Das *zweite* Geschäft besteht in der *Bildung der Normalgleichungen*, die auf die Berechnung des Summensystems

$$\begin{matrix} [aa] & [ab] & [ac] & [al] \\ & [bb] & [bc] & [bl] \\ & & [cc] & [cl] \end{matrix} \tag{2}$$

hinauskommt und bei zahlreichen Beobachtungen und vielziffrigen Koeffizienten den beschwerlichsten Teil der Arbeit bildet. Zu ihrer Erleichterung können Quadrattafeln[1]) und Produkttafeln[2]) gute Dienste leisten. In Ermangelung der letzteren können die nichtquadratischen Summen wie $[ab]$, $[ac]$ usw. mit einer Mehrarbeit auch mittels Quadrattafeln berechnet werden, weil

$$[ab] = \frac{[(a+b)^2] - [aa] - [bb]}{2}$$

ist.[3])

---

1) Solche finden sich in vielen Logarithmenwerken.

2) Rechentafeln von A. L. Crelle (5. Ausg. 1880) und H. Zimmermann, 1889. — Für die Bildung von Quadraten und Produkten sind auch Tafeln der Viertelquadrate geeignet; die ausgedehnteste und verläßlichste derartige Tafel ist die von J. Blater, Wien 1887

3) Bei Beobachtungen ungleicher Genauigkeit kommt noch die Multiplikation mit den Gewichten hinzu. Das zu berechnende Summensystem lautet hier:

$$\begin{matrix} [paa] & [pab] & [pac] & [pal] \\ & [pbb] & [pbc] & [pbl] \\ & & [pcc] & [pcl], \end{matrix}$$

Das *dritte* Geschäft ist die *Auflösung der Normalgleichungen* und in Verbindung damit die Gewichtsbestimmung. Die Aufstellung der reduzierten Normalgleichungen, welche hierzu erforderlich ist, bedingt die Bildung der Koeffizienten

$$\begin{array}{ccc} [bb \cdot 1] & [bc \cdot 1] & [bl \cdot 1] \\ & [cc \cdot 2] & [cl \cdot 2] \end{array} \tag{3}$$

nach den Schematen der Nr. 161; sind diese berechnet, so kann nach Ausführung einiger Divisionen das System:

$$\begin{aligned} x + \frac{[ab]}{[ab]} y + \frac{[ac]}{[aa]} z &= \frac{[al]}{[aa]} \\ y + \frac{[bc \cdot 1]}{[bb \cdot 1]} z &= \frac{[bl \cdot 1]}{[bb \cdot 1]} \\ z &= \frac{[cl \cdot 2]}{[cc \cdot 2]} \end{aligned} \tag{4}$$

ziffermäßig hergestellt werden; aus ihm erhält man (durch sukzessive Substitution) z, y, x. Gleichzeitig ist das Gewicht von z bestimmt; über die Ermittelung der Gewichte von y und z ist in Nr. 161 das Nähere angegeben.

Das *vierte* Geschäft betrifft die Berechnung der *scheinbaren Fehler*. Man erhält sie durch Einsetzung von x, y, z in die Fehlergleichungen (1).

Das *letzte* Geschäft ist die *Genauigkeitsbestimmung*. Sie erfordert die Berechnung des mittleren Fehlers der einzelnen Beobachtung oder der Gewichtseinheit. Dazu ist die Bildung von $[\lambda\lambda]$, beziehungsweise $[p\lambda\lambda]$ nötig, wozu wieder Quadrat- und eventuell auch Produkttafeln verwendet werden können. Hierauf hat man bei u Unbekannten

$$\mu = \sqrt{\frac{[\lambda\lambda]}{n-u}}, \quad \text{beziehungsweise} = \sqrt{\frac{[p\lambda\lambda]}{n-u}}. \tag{5}$$

Daraus berechnen sich mittels der Gewichte p_x, p_y, p_z der Unbekannten deren mittlere Fehler:

$$\mu_x = \frac{\mu}{\sqrt{p_x}}, \quad \mu_y = \frac{\mu}{\sqrt{p_y}}, \quad \mu_z = \frac{\mu}{\sqrt{p_z}}. \tag{6}$$

Bei der Ausführung umfangreicher Rechnungen ist die Anwendung von *Kontrollen* unbedingt erforderlich. Die Theorie gibt deren einige

---

ist aber bei den weiteren Rechenprozessen genau so wie (2) zu behandeln, wird daher häufig auch in der Form

$$\begin{array}{cccc} (aa) & (ab) & (ac) & (al) \\ & (bb) & (bc) & (bl) \\ & & (cc) & (cl) \end{array}$$

geschrieben. Dementsprechend werden auch weiter unten, in (3), (4), . . . die eckigen Klammern durch runde ersetzt.

unmittelbar an die Hand. So besteht eine durchgreifende Probe nach erfolgter Auflösung der Normalgleichungen und Berechnung der λ darin, daß man prüft, ob

$$[a\lambda] = 0, \quad [b\lambda] = 0, \quad [c\lambda] = 0 \tag{7}$$

sei, wie es die Theorie verlangt [s. Nr. 157, Gleichungen (14)]. Bei ungleich genauen Beobachtungen treten an Stelle von (7) die Gleichungen:

$$[pa\lambda] = 0, \quad [pb\lambda] = 0, \quad [pc\lambda] = 0 \tag{7*}$$

[s. Nr. 160, Gleichungen (6*)].

Eine andere, ebenfalls durchgreifende Kontrolle besteht in einer zweiten Berechnung von $[\lambda\lambda]$, die sich wie folgt ergibt. Aus

$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z$$

folgt, wenn man quadriert und addiert,

$$\begin{aligned}[\lambda\lambda] = [ll] &+ \{[aa]x + [ab]y + [ac]z - 2[al]\}x \\ &+ \{[ab]x + [bb]y + [bc]z - 2[bl]\}y \\ &+ \{[ac]x + [bc]y + [cc]z - 2[cl]\}z,\end{aligned}$$

und dies reduziert sich vermöge der Normalgleichungen auf

$$[\lambda\lambda] = [ll] - [al]x - [bl]y - [cl]z.$$

Um sich nun von den berechneten Werten der Unbekannten unabhängig zu machen, schreibe man die erste der Gleichungen (4) darunter in der Anordnung

$$0 = -\frac{[al]}{[aa]} + x + \frac{[ab]}{[aa]}y + \frac{[ac]}{[aa]}z,$$

multipliziere sie mit $[al]$ und addiere sie zur vorigen; dadurch entsteht:

$$[\lambda\lambda] = [ll] - \frac{[al]^2}{[aa]} - [bl\cdot 1]y - [cl\cdot 1]z.$$

Hierunter setze man die zweite der Gleichungen (4):

$$0 = -\frac{[bl\cdot 1]}{[bb\cdot 1]} \qquad + y + \frac{[bc\cdot 1]}{[bb\cdot 1]}z,$$

multipliziere sie mit $[bl\cdot 1]$ und addiere zur vorigen; es entsteht:

$$[\lambda\lambda] = [ll] - \frac{[al]^2}{[aa]} - \frac{[bl\cdot 1]^2}{[bb\cdot 1]} - [cl\cdot 2]z.$$

Die letzte der Gleichungen (4) nach Multiplikation mit $[cl\cdot 2]$ dazugefügt, erhält man schließlich:

$$[\lambda\lambda] = [ll] - \frac{[al]^2}{[aa]} - \frac{[bl\cdot 1]^2}{[bb\cdot 1]} - \frac{[cl\cdot 2]^2}{[cc\cdot 2]}. \tag{8}$$

Hat man dem Schema (2) gleich von Anfang an die Summe $[ll]$ angereiht, so erfordert die Durchführung dieser Formel (8) nur noch Größen, welche schon wegen der Bildung der reduzierten Normal-

gleichungen (4) notwendig sind. Übrigens ist die rechte Seite von (8) gleichbedeutend mit dem Symbol $[ll \cdot 3]$ und kann daher auch nach den Schematen von Nr. 161 gebildet werden; denn es ist

$$[ll] - \frac{[al]^2}{[aa]} = [ll \cdot 1]$$

$$[ll] - \frac{[al]^2}{[aa]} - \frac{[bl \cdot 1]^2}{[bb \cdot 1]} = [ll \cdot 1] - \frac{[bl \cdot 1]^2}{[bb \cdot 1]} = [ll \cdot 2],$$

daher

$$[\lambda\lambda] = [ll \cdot 2] - \frac{[cl \cdot 2]^2}{[cc \cdot 2]} = [ll \cdot 3],$$

wenn man das Gesetz der Schemata fortbildet.

Bei großen Ausgleichungsarbeiten führt man eine die ganze Rechnung begleitende Kontrolle durch, auf die hier nicht eingegangen werden soll.[1])

164. Bedeutung der Resultate. Ableitung empirischer Formeln. Die Theorie vermittelnder Beobachtungen, wie sie hier dargestellt worden ist, beruht auf gewissen Voraussetzungen. Halten wir an dem Falle fest, daß es sich um eine lineare Beziehung handle, so wird zunächst angenommen, daß in der Gleichung

$$\varepsilon_\nu = -l_\nu + a_\nu X + b_\nu Y + c_\nu Z \tag{1}$$

ε_ν den Fehler der Beobachtung l_ν vorstelle und die Parameter a_ν, b_ν, c_ν fehlerfrei bestimmt seien. Weiter wird vorausgesetzt, daß ε_ν dem Gesetze

$$\frac{h_\nu}{\sqrt{\pi}} e^{-h_\nu^2 \varepsilon_\nu^2}$$

folge. Unter diesen Voraussetzungen hat sich die Methode der kleinsten Quadrate als diejenige lineare Kombination der Gleichungen (1) ergeben, welche die mit dem kleinsten Fehlerrisiko verbundenen Werte der Unbekannten liefert, wobei der Begriff des Fehlerrisikos in der Fassung von Nr. 140 gedacht ist.

Unter minder einschränkenden Voraussetzungen, nämlich unter der Annahme, daß die ε_ν von konstanten Anteilen frei seien, daß also das Gesetz, welchem das einzelne ε_ν folgt, irgend eine gerade Funktion von ε_ν sei, hat Gauß[2]) die Methode der kleinsten Quadrate als diejenige lineare Kombination der Gleichungen (1) nachgewiesen, welche Werte der Unbekannten mit den kleinsten mittleren Fehlern liefert. Auch hier ist das Prinzip des kleinsten Fehlerrisikos zur Grundlage genommen, jedoch in dem besonderen Sinne, daß hierunter die Quadratwurzel aus dem mittleren Fehlerquadrat verstanden wird.

---

1) Näheres hierüber sowie über die planmäßige Durchführung der Arbeit unter Verwendung verschiedener Rechenmittel findet man in Helmert, Die Ausgleichungsrechnung etc., 2. Auflage, Leipzig 1907, p. 131—136, 148—171.

2) Theoria combinat.

Kann ferner angenommen werden, daß die scheinbaren Fehler

$$\lambda_\nu = -l_\nu + a_\nu x + b_\nu y + c_\nu z, \tag{2}$$

welche aus der Annahme x, y, z für X, Y, Z resultieren, dem Gesetz $\frac{h_\nu}{\sqrt{\pi}} e^{-h_\nu^2 \lambda_\nu^2}$ folgen, so erweist sich die Methode der kleinsten Quadrate als dasjenige Verfahren, das die wahrscheinlichsten Werte der Unbekannten liefert, weil sie der Koexistenz der Werte λ_ν ($\nu = 1, 2, \cdots n$) die größte Wahrscheinlichkeit verleiht.

Die Methode der kleinsten Quadrate wird aber auch in Fällen zur Anwendung gebracht, wo die Sachlage eine von der obigen verschiedene ist. Man denke zunächst an den Fall, daß nicht bloß l_ν, sondern daß auch die Parameterwerte a_ν, b_ν, c_ν Resultate der Beobachtung und daher Fehlern unterworfen seien, dann hat ε_ν die Bedeutung des Fehlers der Funktion $-l_\nu + a_\nu X + b_\nu Y + c_\nu Z$. Wüßte man, daß die Fehler von l_ν, a_ν, b_ν, c_ν sämtlich Gesetze von der Form $\frac{h}{\sqrt{\pi}} e^{-h^2 \delta^2}$ befolgen, so könnte nach Nr. 138 das Gesetz bestimmt werden, nach dem ε_ν sich richtet, und damit wäre auch die der Gleichung (1) zuzuschreibende Präzision gefunden, alles dies jedoch unter der Voraussetzung, daß man gute Näherungswerte von X, Y, Z im voraus kennt.

Ein weites und wichtiges Gebiet der Anwendung der Methode der kleinsten Quadrate bildet die *Ableitung empirischer Formeln.*

Es liege ein Beobachtungsmaterial vor, das sich auf eine Größe V und auf Parameter a, b, c bezieht, von welchen jene abhängt; dasselbe bestehe in n Reihen zusammengehöriger Werte

$$l_\nu;\quad a_\nu,\quad b_\nu,\quad c_\nu \qquad (\nu = 1, 2, \cdots n)$$

Man soll dieses Beobachtungsmaterial durch eine analytische Gleichung

$$V = F(X, Y, Z) \tag{3}$$

darstellen, die entweder aus theoretischen Erwägungen hervorgegangen ist oder als Hypothese über den Zusammenhang zwischen V, a, b, c angenommen wurde und außer diesen Größen noch u ($= 3$) unbekannte Konstanten X, Y, Z enthält.

Man schlägt nun zur Lösung dieser Aufgabe den Weg ein, den die Methode der kleinsten Quadrate für dieses Problem, als einen Fall der Ausgleichungsrechnung, vorschreibt, bestimmt die „vorteilhaftesten" Werte x, y, z der Unbekannten und bildet mit diesen die empirische Formel

$$v = F(x, y, z), \tag{4}$$

welche gestattet, zu jeder Wertgruppe der Parameter a, b, c den zugehörigen Wert von V zu *berechnen.*

In solchen Fällen handelt es sich nicht mehr um die „Ausgleichung von *Beobachtungsfehlern*“; denn die λ_ν, welche sich durch die Differenzen

$$\lambda_\nu = - l_\nu + v_\nu$$

darstellen, stammen jetzt nicht allein von den Fehlern her, welche den beobachteten Größen l_ν, a_ν, b_ν, c_ν anhaften, sondern und meist vielmehr aus den vereinfachenden Annahmen, welche bei der theoretischen Ableitung des Ansatzes (3) gemacht worden sind, beziehungsweise aus der Hypothese, deren analytischen Ausdruck dieser Ansatz bildet. Die gewonnene Formel (4) kann als eine solche bezeichnet werden, der das Beobachtungsmaterial in seiner Gänze „möglichst genau genügt“, und die Summe $[\lambda\lambda]$, die kleinste unter allen, so lange man an der Form (3) festhält, repräsentiert ein „Maß für das Genügen“. Die Gewichte der Unbekannten, wenn man sie mitbestimmt, belehren über den Grad der Sicherheit, mit welchem jede einzelne Unbekannte bestimmt ist. Hat man über dasselbe Beobachtungsmaterial zwei verschiedene Hypothesen in Form von Gleichungen aufgestellt und zu jeder Hypothese die empirische Formel bestimmt, so zeigen die zugehörigen Summen $[\lambda\lambda]$, welche der Hypothesen das Beobachtungsmaterial in seiner Gänze besser darzustellen geeignet ist; man wird von diesem Gesichtspunkte aus jener den Vorzug geben, der das kleinere $[\lambda\lambda]$ zukommt.[1])

Eine allgemeine Entscheidung darüber, wann eine empirische Formel überhaupt geeignet ist, einen Komplex von Beobachtungen analytisch darzustellen, kann nicht gefällt werden. Es hängt dies davon ab, ob die λ_ν, zu welchen die Formel führt, innerhalb der Fehlergrenze liegen, mit welcher V beobachtet werden kann, oder ob sie Abweichungen vorstellen, die mit Rücksicht auf den verfolgten Zweck vernachlässigt werden dürfen.

165. Beispiel LX. Die Linien des sogenannten Wasserstoffspektrums zeigen eine von Balmer entdeckte Gesetzmäßigkeit; man kann nämlich die Wellenlänge der aufeinanderfolgenden Linien H_α, H_β, H_γ, $\cdots$ des Spektrums durch die Formel

$$L = \frac{m^2}{m^2 - 4} X$$

ausdrücken, in welcher m die aufeinanderfolgenden ganzen Zahlen von 3 an und X einen konstanten Faktor bedeutet; zur Berechnung

1) Über diese erweiterte Auffassung der Methode der kleinsten Quadrate als eines Verfahrens zur Lösung von Problemen des „möglichst genau Genügens“, des „möglichst nahe Liegens“ (bei geometrischer Auslegung) vergleiche man die kleine Schrift von R. Henke, Über die Methode der kleinsten Quadrate, 2. Aufl., Leipzig 1894.

dieses Faktors sollen die Ergebnisse der von Ames durchgeführten Messungen der Wellenlängen benützt werden, deren Resultate in der weiter unten aufgestellten Tabelle aus der ersten und vierten Kolonne ersichtlich sind.[1])

Bezeichnet man den a priori bekannten Parameter $\frac{m^2}{m^2-4}$ mit a, die beobachtete Wellenlänge mit l, so gibt die obige Formel Anlaß zu einem System von Fehlergleichungen der Form:

$$\lambda = -l + ax.$$

Bestimmt man für X aus irgend einer der Beobachtungen den auf Einheiten abgerundeten Näherungswert X_0, so ergibt sich

$$X_0 = 3647;$$

nennt man den vorteilhaftesten Wert der hieran anzubringenden Korrektur ξ, so nehmen die Fehlergleichungen die Gestalt

$$\lambda = -l' + a\xi$$

an, wobei

$$l' = l - aX_0.$$

Die einzige Normalgleichung

$$[aa]\xi = [al']$$

liefert

$$\xi = \frac{[al']}{[aa]} = \frac{2,80}{17,47} = 0,16;$$

folglich ist

$$x = X_0 + \xi = 3647,16,$$

und die empirische Formel zur Bestimmung der Wellenlänge lautet:

$$L = 3647,16 \frac{m^2}{m^2-4}.$$

Linie	m	a	l	l'	aa	al'	$l-L$
$H\alpha$	3	1,8000	6564,97	0,37	3,2400	0,67	0,08
β	4	1,3333	4862,93	0,39	1,7769	0,52	0,06
γ	5	1,1905	4342,00	0,25	1,4185	0,48	0,07
δ	6	1,1250	4103,11	0,33	1,2656	0,37	0,05
ε	7	1,0889	3971,40	0,18	1,1859	0,19	0,02
ζ	8	1,0667	3890,30	0,15	1,1385	0,16	0,02
η	9	1,0519	3836,80	0,34	1,1067	0,36	0,17
ϑ	10	1,0417	3799,20	0,12	1,0858	0,12	0,09
ι	11	1,0342	3771,90	0,17	1,0692	0,17	0,01
$\varkappa$	12	1,0286	3751,30	0,00	1,0588	0,00	— 0,01
λ	13	1,0242	3735,30	0,04	1,0486	0,04	— 0,27
μ	14	1,0208	3722,80	— 0,17	1,0424	— 0,17	— 0,33
ν	15	1,0181	3712,90	— 0,11	1,0363	— 0,11	— 0,27
					17,4732	2,80	

1) W. Nernst, Theoretische Chemie, 2. Aufl., 1898, p. 195—196.

Die vorstehende Tabelle gibt außer den Beobachtungen auch die zur Berechnung von x benützten Größen sowie die Differenzen $l - L$, d. i. Beobachtung — Berechnung. Eine Genauigkeitsbestimmung ist nicht vorgenommen worden.

166. Beispiel LXI. *Ausgleichung von Höhenmessungen.* Es sind die Höhenabstände folgender Punktepaare gemessen worden:

$$\begin{aligned}(AB) = l_1,\quad (AC) = l_2,\quad & (AD) = l_3\\ (BC) = l_4,\quad & (BD) = l_5\\ & (CD) = l_6.\end{aligned}$$

Zur Festlegung der relativen Höhenlage der vier Punkte A, B, C, D genügen aber drei Höhenabstände, somit sind drei Beobachtungen überzählig. Faßt man die drei ersten Höhenabstände als die Unbekannten X, Y, Z auf, so ergeben sich folgende Fehlergleichungen:

$$\begin{aligned}\lambda_1 &= -l_1 + x\\ \lambda_2 &= -l_2 \quad + y\\ \lambda_3 &= -l_3 \quad\quad + z\\ \lambda_4 &= -l_4 - x + y\\ \lambda_5 &= -l_5 - x \quad + z\\ \lambda_6 &= -l_6 \quad - y + z.\end{aligned}$$

Sind alle Messungsergebnisse als gleich genau zu betrachten, so hat man zur Bestimmung von x, y, z die Normalgleichungen:

$$\begin{aligned}3x - y - z &= l_1 - l_4 - l_5\\ -x + 3y - z &= l_2 + l_4 - l_6\\ -x - y + 3z &= l_3 + l_5 + l_6.\end{aligned}$$

Zu ihrer Lösung bilde man die Summe:

$$x + y + z = l_1 + l_2 + l_3$$

und addiere sie folgeweise zur ersten, zweiten und dritten Gleichung; dadurch erhält man:

$$\begin{aligned}x &= \frac{1}{4}\{2l_1 + l_2 + l_3 - l_4 - l_5\}\\ y &= \frac{1}{4}\{l_1 + 2l_2 + l_3 + l_4 - l_6\}\\ z &= \frac{1}{4}\{l_1 + l_2 + 2l_3 + l_5 + l_6\}.\end{aligned}$$

Für $[\lambda\lambda]$ ergibt sich nach der Formel

$$[\lambda\lambda] = [ll] - [al]x - [bl]y - [cl]z$$

mit Benützung der rechten Seiten der Normalgleichungen der Ausdruck:

$$[\lambda\lambda] = \frac{1}{2}\{[ll] - l_1 l_2 - l_1 l_3 + l_1 l_4 + l_1 l_5 - l_2 l_3 - l_2 l_4 \\ + l_2 l_6 - l_3 l_5 - l_3 l_6 - l_4 l_5 + l_4 l_6 - l_5 l_6\};$$

die Quadratwurzel aus dem dritten Teile dieses Wertes ist der mittlere Fehler einer einzelnen Beobachtung.

Die Gewichtsgleichungen für x sind:

$$\begin{aligned} 3Q_{11} - Q_{12} - Q_{13} &= 1 \\ -Q_{11} + 3Q_{12} - Q_{13} &= 0 \\ -Q_{11} - Q_{12} + 3Q_{13} &= 0; \end{aligned}$$

verfährt man mit ihnen wie mit den Normalgleichungen, so ergibt sich:

$$Q_{11} = \frac{1}{2}; \text{ ebenso ist } Q_{22} = \frac{1}{2}, \quad Q_{33} = \frac{1}{2}.$$

Es ist hiernach der mittlere Fehler jeder Unbekannten dem $\sqrt{\frac{1}{2}}$-fachen mittleren Fehler einer Beobachtung gleich.

167. Beispiel LXII. *Stationsausgleichung.* Bei dieser handelt es sich um eine ähnliche Aufgabe wie bei der Ausgleichung von Höhenmessungen.

Von einem Punkte L aus sind folgende Winkel zwischen den Punkten E, S, A, G, Fig. 17, gemessen worden:

$$\begin{aligned} (ES) &= l_1, \quad (EA) = l_2, \quad (EG) = l_3 \\ (SA) &= l_4, \quad (SG) = l_5. \end{aligned}$$

Die nachfolgend angegebenen Ergebnisse sind Mittelwerte aus der daneben angeführten Anzahl einfacher Beobachtungen:

$$\begin{aligned} l_1 &= 59^0 53' 44'',03, \quad 16 \\ l_2 &= 98\ 53\ 23\ ,41, \quad 8 \\ l_3 &= 271\ 17\ 29\ ,38, \quad 8 \\ l_4 &= 38\ 59\ 38\ ,97, \quad 8 \\ l_5 &= 211\ 23\ 46\ ,38, \quad 8 \end{aligned}$$ [1]).

E S G L A

Fig. 17.

Da zur relativen Bestimmung der vier Richtungen LE, LS, LA, LG drei Winkel ausreichen, so sind zwei Messungen überschüssig. Wählt man

$$X = (ES), \quad Y = (EA), \quad Z = (EG)$$

1) Die Europäische Längengradmessung. I. Heft. Veröffentl. d. Königl. Preuß. Geod. Inst. etc. Berlin 1893, p. 205.

als die Unbekannten, die rohen Beobachtungswerte l_1, l_2, l_3 als deren Näherungswerte, so ergeben sich für die vorteilhaftesten Verbesserungen x, y, z an den Näherungswerten die Fehlergleichungen:

$$\begin{array}{ll|l}
\lambda_1 = & \quad x & \text{Gewicht } 2 \\
\lambda_2 = & \qquad\quad y & \quad " \quad 1 \\
\lambda_3 = & \qquad\qquad\quad z & \quad " \quad 1 \\
\lambda_4 = & 0{,}41 - x + y & \quad " \quad 1 \\
\lambda_5 = & -1{,}03 - x \quad + z & \quad " \quad 1
\end{array}$$

0,41 entsteht als Differenz aus $l_2 - l_1$ und l_4, $-1{,}03$ als Differenz aus $l_3 - l_1$ und l_5.

Die Bildung der Normalgleichungen gestaltet sich sehr einfach; es ist nämlich

$$\begin{array}{llll}
[paa] = 4, & [pab] = -1, & [pac] = -1, & [pal] = -0{,}62 \\
 & [pbb] = \;\; 2, & [pbc] = \;\; 0, & [pbl] = -0{,}41 \\
 & & [pcc] = \;\; 2, & [pcl] = \;\; 1{,}03;
\end{array}$$

sie lauten daher:

$$\begin{aligned}
4x - y - z &= -0{,}62 \\
-x + 2y \qquad &= -0{,}41 \\
-x \qquad + 2z &= \;\; 1{,}03
\end{aligned}$$

und geben ohne Mühe

$$x = -0'',103, \quad y = -0'',257, \quad z = 0'',464.$$

Man berechnet daraus

$$\begin{aligned}
\lambda_1 &= -0{,}103 \\
\lambda_2 &= -0{,}257 \\
\lambda_3 &= \;\; 0{,}464 \\
\lambda_4 &= \;\; 0{,}257 \\
\lambda_5 &= -0{,}464,
\end{aligned}$$

und zwischen den ausgeglichenen Winkeln

$$\begin{aligned}
l_1 + \lambda_1 &= \;\; 59^{\circ}53'43'',927 \\
l_2 + \lambda_2 &= \;\; 98 \;\; 53 \;\; 23 \;\; ,153 \\
l_3 + \lambda_3 &= 271 \;\; 17 \;\; 29 \;\; ,844 \\
l_4 + \lambda_4 &= \;\; 38 \;\; 59 \;\; 39 \;\; ,226 \\
l_5 + \lambda_5 &= 211 \;\; 23 \;\; 45 \;\; ,917
\end{aligned}$$

bestehen keine Widersprüche mehr.

Aus den drei Systemen von Gewichtsgleichungen:

$$\left.\begin{array}{l} 4Q_{11} - Q_{12} - Q_{13} = 1 \\ -Q_{11} + 2Q_{12} \qquad = 0 \\ -Q_{11} \qquad + 2Q_{13} = 0 \end{array}\right\} \quad \left.\begin{array}{l} 4Q_{21} - Q_{22} - Q_{23} = 0 \\ -Q_{21} + 2Q_{22} \qquad = 1 \\ -Q_{21} \qquad + 2Q_{23} = 0 \end{array}\right\}$$

$$\left.\begin{array}{l} 4Q_{31} - Q_{32} - Q_{33} = 0 \\ -Q_{31} + 2Q_{32} \qquad = 0 \\ -Q_{31} \qquad + 2Q_{33} = 1 \end{array}\right\}$$

findet man leicht

$$Q_{11} = \frac{1}{3}, \qquad Q_{22} = \frac{7}{12}, \qquad Q_{33} = \frac{7}{12}.$$

Ferner berechnet sich direkt

$$[p\lambda\lambda] = 0{,}581028,$$

daraus der mittlere Fehler der Gewichtseinheit:

$$\mu = \sqrt{\frac{0{,}581028}{5-3}} = 0'',539;$$

endlich

$$\mu_x = 0'',539\sqrt{\frac{1}{3}} = 0'',309, \qquad \mu_y = 0'',539\sqrt{\frac{7}{12}} = 0'',415 = \mu_z.$$

168. Beispiel LXIII. *Empirische Formel für die Länge des Sekundenpendels.* Die Länge des Sekundenpendels an einem Orte unter der geographischen Breite B läßt sich durch die Formel:

$$L = X - Y\cos 2B$$

darstellen, in welcher X, Y Konstanten bedeuten. Da für $B = 45^0$ $L = X$ wird, so heißt X die auf 45^0 reduzierte Pendellänge, und Y ist die Konstante, durch welche die Reduktion bewirkt wird. Auf Grund des nachfolgend mitgeteilten Beobachtungsmaterials ist die Bestimmung der Konstanten vorzunehmen.[1])

Ist l eine unter der Breite B gemessene Pendellänge, so gibt sie Anlaß zu der Fehlergleichung:

$$\lambda = -l + x - y\cos 2B.$$

Aus einer vorläufigen Rechnung ergab sich für y der Näherungswert 0,002636 in Metern; setzt man

$$y = 0{,}002636\left(1 - \frac{\eta}{100}\right),$$

so daß η die prozentische Verbesserung dieses Wertes bedeutet, so verwandelt sich die Fehlergleichung in:

$$\lambda = -l - 0{,}002636\cos 2B + x + 0{,}00002636\,\eta\cos 2B;$$

---

1) F. R. Helmert, Die mathem. u. physik. Theor. d. höheren Geodäsie. II. Bd., Leipzig 1884, p. 238 ff.

mit

$$l + 0{,}002636 \cos 2B = l', \quad 0{,}00002636 \cos 2B = b$$

schreibt sie sich endlich:

$$\lambda = -l' + x + b\eta.$$

In der nachstehenden Tabelle sind die Werte von l' und b für acht verschiedene Breiten zwischen 0 und 80° und die weiter zur Durchführung der Rechnung erforderlichen Größen zusammengestellt.

Nr.	l' in Mikrons	b in Mikrons	bb	bl'	λ	$\lambda\lambda$
1	993 568	25,5	650,25	25 335 984,0	− 13,8	190,44
2	993 559	22,9	524,41	22 752 501,1	− 5,4	29,16
3	993 528	17,5	306,25	17 386 740,0	+ 24,6	605,16
4	993 552	10,4	108,16	10 332 940,8	− 0,8	0,64
5	993 551	1,5	2,25	1 490 326,5	− 1,4	1,96
6	993 555	− 8,0	64,00	− 7 948 440,0	− 7,2	51,84
7	993 540	− 16,2	262,44	− 16 095 348,0	+ 6,1	37,21
8	993 549	− 22,5	506,25	− 22 354 852,5	− 4,0	16,00
	7 948 402	31,1	2424,01	30 899 851,9		934,41
	$[l']$	$[b]$	$[bb]$	$[bl']$		$[\lambda\lambda]$

Da in dem vorliegenden Falle $a_\nu = 1$ ist für alle Werte von ν, so ist $[aa] = 8$, $[ab] = [b]$, $[al'] = [l']$, und es lauten daher die Normalgleichungen:

$$8x + \quad 31{,}1\eta = \quad 7\,948\,402$$

$$31{,}1x + 2424{,}0\eta = 30\,899\,852\,.$$

Multipliziert man die erste mit $\frac{31{,}1}{8}$ und subtrahiert sie von der zweiten, so entsteht

$$2302{,}7\eta = 447\,,$$

woraus

$$\eta = 0{,}19$$

und

$$p_\eta = 2302{,}7\,.$$

Kehrt man die Normalgleichungen um:

$$2424{,}0\eta + 31{,}1x = 30\,899\,852$$

$$31{,}1\eta + \quad 8x = \quad 7\,948\,402\,,$$

multipliziert die erste mit $\frac{31{,}1}{2424{,}0}$ und subtrahiert sie von der zweiten, so ergibt sich

$$7{,}6x = 7\,551\,596\,,$$

daraus

$$x = 993\,549$$

und

$$p_x = 7{,}6.$$

Mit Hilfe von x, η können nun aus den Fehlergleichungen die λ gerechnet werden; so ist

$$\lambda_1 = -993\,568 + 993\,549 + 25{,}5 \cdot 0{,}19 = -13{,}8$$

usw.; die Werte sind in die Tabelle eingestellt.

Der mittlere Fehler einer Gleichung ist hiernach

$$\mu = \sqrt{\frac{932{,}41}{8-2}} = 12{,}5 \text{ Mikrons},$$

der mittlere Fehler in der Bestimmung von η:

$$\mu_\eta = \frac{12{,}5}{\sqrt{2308}} = 0{,}26.$$

Wie aus den Gewichten zu ersehen, ist die Bestimmung von x weit unsicherer als die von η; der mittlere Fehler von x:

$$\mu_x = \frac{12{,}5}{\sqrt{7{,}6}} = 4{,}5$$

beträgt $4\frac{1}{2}$ Mikrons.

Für y ergibt sich jetzt der Wert

$$y = 0{,}002636\,(1 - 0{,}0019) = 0{,}002631$$

mit dem mittleren Fehler

$$0{,}002631 \cdot 0{,}0026.$$

Mithin lautet die empirische Formel mit Angabe der mittleren Unsicherheit in y:

$$\mathfrak{L} = 993\,549 - 2631\,(1 \mp 0{,}0026) \cos 2B$$

in Mikrons, oder, wenn man $\cos 2B$ durch $1 - 2\sin^2 B$ ersetzt:

$$\mathfrak{L} = 0{,}990918\,\{1 + (0{,}005310 \mp 0{,}000014) \sin^2 B\}$$

in Metern.

169. Beispiel LXIV. *Empirische Formel für die Spannkraft des gesättigten Wasserdampfes.* Magnus hat im Anschlusse an Untersuchungen von August für den Zusammenhang zwischen der Spannkraft S und der Temperatur t des gesättigten Wasserdampfes die Formel

$$S = X\,10^{\frac{tY}{Z+t}}$$

aufgestellt, in welcher X, Y, Z Konstanten bedeuten. Für diese sollen auf Grund des in den zwei ersten Kolonnen der weiter unten folgen-

den Tafel mitgeteilten, von **Magnus** stammenden Beobachtungsmaterials die vorteilhaftesten Werte gefunden werden.

Es bieten sich zur Lösung dieser Aufgabe zwei Wege dar (s. Nr. 162). Logarithmiert man (1), so entsteht:

$$\log S = \log X + \frac{tY}{Z+t};$$

entwickelt man den Bruch mit Hilfe von Näherungswerten Y_0, Z_0 für Y, Z nach den Korrektionen y, z, so gibt er bei Beschränkung auf deren erste Potenzen:

$$\frac{tY_0}{Z_0+t} + \frac{t}{Z_0+t}y - \frac{tY_0}{(Z_0+t)^2}z,$$

und wenn l die zu t gehörige Beobachtung ist, so entsteht die Fehlergleichung:

$$\lambda = -\log l + \frac{tY_0}{Z_0+t} + \log x + \frac{t}{Z_0+t}y - \frac{tY_0}{(Z_0+t)^2}z, \tag{2}$$

die linear ist in bezug auf $\log x$, y, z; bei weiterer Führung der Rechnung wäre ihr das Gewicht

$$\frac{1}{\left(\frac{\partial \log l}{\partial l}\right)^2} = \frac{1}{\left(\frac{M}{l}\right)^2} = \left(\frac{l}{M}\right)^2$$

oder ein zu l^2 proportionales Gewicht beizulegen, weil M eine Konstante, den Modul der gemeinen Logarithmen, bedeutet.

Der zweite Weg besteht darin, daß man die Funktion in (1) selbst mit Hilfe von Näherungswerten X_0, Y_0, Z_0 aller drei Konstanten nach den Korrektionen x, y, z entwickelt. Setzt man

$$\left.\begin{aligned}
X_0\, 10^{\frac{tY_0}{Z_0+t}} &= S_0,\\
\frac{\partial S_0}{\partial X_0} = 10^{\frac{tY_0}{Z_0+t}} = a, \quad \frac{\partial S_0}{\partial Y_0} &= X_0\, 10^{\frac{tY_0}{Z_0+t}} \cdot \frac{tl \cdot 10}{Z_0+t} = \frac{X_0\, tl \cdot 10}{Z_0+t}\, a = b,\\
\frac{\partial S_0}{\partial Z_0} = X_0\, 10^{\frac{tY_0}{Z_0+t}} \cdot \frac{-Y_0\, tl \cdot 10}{(Z_0+t)^2} &= \frac{-Y_0}{Z_0+t}\, b = c,\\
l - S_0 &= l',
\end{aligned}\right\} \tag{3}$$

so ergibt sich für die Beobachtung l die lineare Fehlergleichung:

$$\lambda = -l' + ax + by + cz. \tag{4}$$

Dieser Vorgang ist dem früheren, hauptsächlich weil er die beschwerliche Berücksichtigung der Gewichte erspart, vorzuziehen.

In der folgenden Tabelle sind neben dem Beobachtungsmaterial (t, l) die nach den Formeln (3) gerechneten Koeffizienten der Fehlergleichungen zusammengestellt und die aus der weiteren Rechnung hervorgegangenen Abweichungen λ angefügt. Was die Näherungswerte anlangt, so verschafft man sich solche durch Anwendung der

Gleichung (1) auf drei Beobachtungen; für $t = 0$ hat man sogleich $X_0 = 4{,}53$; die andern zwei Gleichungen bringt man durch Logarithmieren auf die Form:

$$t\,Y_0 - \log \frac{l}{X_0} \cdot Z_0 = t \log \frac{l}{X_0},$$

so daß sie linear werden in bezug auf Y_0, Z_0. Eine derart durchgeführte Rechnung ergab:

$$X_0 = 4{,}53, \quad Y_0 = 7{,}45, \quad Z_0 = 234{,}70.$$

Nr.	Temperatur t^0 C	Spannkraft l mm	a	b	c	l'	λ mm
1	— 5,31	2,95	+ 0,672	— 0,161	+ 0,005	— 0,095	+ 0,036
2	— 3,64	3,45	0,763	— 0,124	+ 0,005	— 0,007	— 0,058
3	+ 0,00	4,525	1,000	0,000	0,000	— 0,005	— 0,078
4	8,01	7,93	1,761	+ 0,605	— 0,018	— 0,049	— 0,082
5	11,98	9,88	2,300	1,165	0,035	— 0,541	+ 0,376
6	16,82	13,52	3,149	2,196	0,064	— 0,746	+ 0,537
7	23,85	22,24	4,867	4,681	0,136	+ 0,194	— 0,490
8	35,95	48,96	9,763	13,523	0,373	— 0,265	— 0,230
9	44,90	71,20	15,717	26,326	0,726	+ 0,002	— 0,697
10	52,12	101,40	22,583	42,806	1,112	— 0,903	+ 0,028
11	58,68	139,72	30,910	64,490	1,636	— 0,893	— 0,165
12	74,47	281,55	62,300	156,520	3,728	— 0,649	— 0,839
13	78,83	330,58	73,152	190,924	4,548	— 1,248	— 0,331
14	82,25	387,56	85,765	232,136	5,455	— 1,365	— 0,288
15	86,21	453,31	100,319	281,098	6,528	— 3,807	+ 2,109
16	91,34	552,20	122,213	357,103	8,160	— 0,592	— 1,115
17	93,66	602,53	133,354	396,752	9,003	— 2,314	+ 0,624
18	99,39	743,49	164,564	510,637	11,381	— 1,916	+ 0,349
19	100,87	784,07	173,547	544,065	12,079	— 6,439	+ 4,922
20	104,64	895,83	198,293	637,758	13,999	+ 3,435	— 4,772

Nach Ausrechnung der Summenkoeffizienten ergeben sich die Normalgleichungen:

$$157\,743{,}1225\,x + 476\,833{,}1814\,y - 10\,722{,}0171\,z = -1816{,}8964$$
$$476\,833{,}1814\,x + 1\,449\,178{,}8203\,y - 32\,531{,}7187\,z = -5302{,}7052$$
$$-10\,722{,}0171\,x - 32\,531{,}7187\,y + 740{,}5679\,z = 121{,}5599.$$

Eliminiert man nach dem Nr. 161 erläuterten Vorgange nach und nach x und y, so gibt die Endgleichung:

$$9{,}9061\,z = 1{,}0019$$

gleichzeitig

$$z = 0{,}1011 \text{ und } p_z = 9{,}91;$$

aus dem auf diese Weise entspringenden System reduzierter Normalgleichungen:

$$x + 3{,}0228\,y - 0{,}0679\,z = -0{,}0115$$
$$y - 0{,}0155\,z = 0{,}0243$$
$$z = 0{,}1011$$

folgt durch Substitution:

$$y = 0{,}0259, \quad x = -0{,}0829.$$

Durch Umstellung der Normalgleichungen ergab sich weiter:

$$p_y = 6548{,}36, \quad p_x = 816{,}35.$$

Hiernach sind die verbesserten Werte der Konstanten:

$$X_0 + x = 4{,}53 - 0{,}0829 = 4{,}4471$$
$$Y_0 + y = 7{,}45 + 0{,}0259 = 7{,}4759$$
$$Z_0 + z = 234{,}70 + 0{,}1011 = 234{,}8011,$$

und es lautet die empirische Formel:

$$S = 4{,}4471 \cdot 10^{\frac{7{,}4759\, t}{234{,}8011 + t}}. \tag{5}$$

Aus

$$[\lambda\lambda] = 55{,}3481$$

berechnet sich der mittlere Fehler einer Gleichung:

$$\mu = \sqrt{\frac{55{,}3481}{20 - 8}} = 1{,}80\,\text{mm};$$

daraus folgen die mittleren Fehler der Konstanten:

$$\mu_x = \frac{1{,}80}{\sqrt{816{,}35}} = 0{,}0630, \quad \mu_y = \frac{1{,}80}{\sqrt{6548{,}36}} = 0{,}0222, \quad \mu_z = \frac{1{,}80}{\sqrt{9{,}91}} = 0{,}5714.$$

Nachstehend ist für einige Temperaturen der nach Formel (5) berechnete Wert von S neben denjenigen gestellt, welcher sich in Kohlrauschs „Leitfaden der praktischen Physik", 8. Aufl., 1896, S. 469—470 angegeben findet. Man wird daraus entnehmen, daß die Formel namentlich für nicht zu hohe Temperaturen den Gang der experimentell bestimmten Werte gut wiedergibt.

Temperatur	Spannung	
	nach (5)	nach Kohlrausch
— 10°	2,1 mm	2,2 mm
0	4,4	4,6
+ 10	9,0	9,1
20	17,2	17,4
30	31,3	31,5
40	54,5	54,9
50	91,3	92,0
60	148,8	148,9
70	231,7	233,3
80	353,1	355,4
90	524,3	525,9
100	760,2	760,0

§ 3. Bedingte Beobachtungen.

170. Problemstellung. Zurückführung auf vermittelnde Beobachtungen. Für die Größen $X_1, X_2, \cdots X_n$ seien durch Beobachtung die Werte $l_1, l_2, \cdots l_n$ erhalten worden. Sind jene *unabhängig* von einander, so kann der Fall unter die Form vermittelnder Beobachtungen gestellt werden wie folgt: Es ist die Größe

$$V = aX_1 + bX_2 + \cdots + gX_n$$

n-mal beobachtet worden, und den einzelnen Beobachtungen entsprechen die nachstehenden Wertgruppen von V und der Parameter $a, b, \cdots g$:

$$\begin{array}{l} l_1;\ 1,\ 0,\ 0,\ \cdots\ 0 \\ l_2;\ 0,\ 1,\ 0,\ \cdots\ 0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ l_n;\ 0,\ 0,\ 0,\ \cdots\ 1. \end{array}$$

Daraus ergibt sich das System der Fehlergleichungen:

$$\begin{aligned} \lambda_1 &= -l_1 + x_1 \\ \lambda_2 &= -l_2 \qquad + x_2 \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \lambda_n &= -l_n \qquad\qquad + x_n, \end{aligned}$$

aus dem wieder das System der Normalgleichungen hervorgeht:

$$\begin{aligned} x_1 &= l_1 \\ x_2 &= l_2 \\ &\cdot\ \cdot\ \cdot \\ x_n &= l_n, \end{aligned}$$

das nichts anderes besagt, als daß in diesem Falle der vorteilhafteste Wert einer jeden Unbekannten der durch die Beobachtung für sie ermittelte Wert sei. Die Summe $[\lambda\lambda]$ nimmt dabei den absolut kleinsten Wert, 0, an. Daraus ist aber nicht auf Fehlerfreiheit der Beobachtungen zu schließen; vielmehr ist, da sich keine Widersprüche ergeben können, kein Mittel zur Beurteilung der Genauigkeit vorhanden.

Anders gestaltet sich die Sache, wenn zwischen den Größen $X_1, X_2, \cdots X_n$ Gleichungen bestehen, welche a priori gegeben sind. Dann muß den vorteilhaftesten Werten $x_1, x_2, \cdots x_n$ die *Bedingung* auferlegt werden, daß auch sie jene Gleichungen erfüllen. Dies hat, da die rohen Beobachtungsergebnisse vermöge ihrer Fehlerhaftigkeit diese Bedingung nicht erfüllen werden, Gleichungen zwischen den Verbesserungen $\lambda_1, \lambda_2, \cdots \lambda_n$ zur Folge, die an $l_1, l_2, \cdots l_n$ angebracht werden müssen, um sie in $x_1, x_2, \cdots x_n$ überzuführen.

Es seien

$$\left.\begin{array}{l} f_1(X_1,\ X_2,\ \cdots\ X_n) = 0 \\ f_2(X_1,\ X_2,\ \cdots\ X_n) = 0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ f_r(X_1,\ X_2,\ \cdots\ X_n) = 0 \end{array}\right\} \qquad (1)$$

die Gleichungen, welchen X_1, X_2, $\cdots$ X_n zu genügen haben — *die Bedingungsgleichungen*. Ihre Anzahl r ist notwendig kleiner als n.

Dann wird verlangt, daß auch

$$\begin{array}{l} f_1(x_1,\ x_2,\ \cdots\ x_n) = f_1(l_1 + \lambda_1,\ l_2 + \lambda_2,\ \cdots\ l_n + \lambda_n) = 0 \\ f_2(x_1,\ x_2,\ \cdots\ x_n) = f_2(l_1 + \lambda_1,\ l_2 + \lambda_2,\ \cdots\ l_n + \lambda_n) = 0 \\ \cdot \\ f_r(x_1,\ x_2,\ \cdots\ x_n) = f_r(l_1 + \lambda_1,\ l_2 + \lambda_2,\ \cdots\ l_n + \lambda_n) = 0 \end{array}$$

sei. Entwickelt man die linken Seiten und setzt dabei zur Abkürzung:

$$\left.\begin{array}{l} f_1(l_1,\ l_2,\ \cdots\ l_n) = w_1 \\ f_2(l_1,\ l_2,\ \cdots\ l_n) = w_2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ f_r(l_1,\ l_2,\ \cdots\ l_n) = w_r \end{array}\right\} \qquad (2)$$

$$\frac{\partial f_1}{\partial l_\nu} = a_\nu, \quad \frac{\partial f_2}{\partial l_\nu} = b_\nu, \quad \cdots \quad \frac{\partial f_r}{\partial l_\nu} = g_\nu, \qquad (3)$$

so entstehen die Gleichungen:

$$\left.\begin{array}{lll} a_1\lambda_1 + a_2\lambda_2 + \cdots + a_n\lambda_n + w_1 = 0 & \text{oder} & [a\lambda] + w_1 = 0 \\ b_1\lambda_1 + b_2\lambda_2 + \cdots + b_n\lambda_n + w_2 = 0 & \text{"} & [b\lambda] + w_2 = 0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ g_1\lambda_1 + g_2\lambda_2 + \cdots + g_n\lambda_n + w_r = 0 & \text{"} & [g\lambda] + w_r = 0. \end{array}\right\} \qquad (4)$$

Diese Form der Bedingungsgleichungen ist streng, wenn die (1) linear waren: im andern Falle stellt sie nur eine Näherung dar, die um so zutreffender ist, je kleiner die λ sind.

Die Größen w_1, w_2, $\cdots$ w_r bedeuten, wie aus ihrer Definition (2) zu ersehen, die *Widersprüche*, die sich ergeben, wenn man in die Bedingungsgleichungen (1) statt der wahren Werte die Beobachtungsergebnisse einsetzt.

Die Gleichungen (4) gestatten, r von den λ, etwa λ_1, λ_2, $\cdots$ λ_r, durch die übrigen *linear* auszudrücken. Ist beispielsweise $r = 4$ und $n = 7$, und setzt man

$$\lambda_5 = \xi, \quad \lambda_6 = \eta, \quad \lambda_7 = \zeta,$$

so wird

$$\begin{aligned}
\lambda_1 &= -\mathfrak{l}_1 + \mathfrak{a}_1\xi + \mathfrak{b}_1\eta + \mathfrak{c}_1\zeta \\
\lambda_2 &= -\mathfrak{l}_2 + \mathfrak{a}_2\xi + \mathfrak{b}_2\eta + \mathfrak{c}_2\zeta \\
\lambda_3 &= -\mathfrak{l}_3 + \mathfrak{a}_3\xi + \mathfrak{b}_3\eta + \mathfrak{c}_3\zeta \\
\lambda_4 &= -\mathfrak{l}_4 + \mathfrak{a}_4\xi + \mathfrak{b}_4\eta + \mathfrak{c}_4\zeta \\
\lambda_5 &= \xi \\
\lambda_6 &= \eta \\
\lambda_7 &= \zeta.
\end{aligned} \tag{5}$$

Das ist aber ein System von Fehlergleichungen, wie es bei vermittelnden Beobachtungen vorlag. Hat man auf Grund des Prinzips „$[\lambda\lambda]$ ein Minimum" die vorteilhaftesten Werte von ξ, η, ζ berechnet und mittels derselben $\lambda_1, \lambda_2, \cdots \lambda_7$ bestimmt, so sind auch $x_1, x_2, \cdots x_7$ gefunden.

Es ist demnach die Ausgleichung von *n Beobachtungen mit r Bedingungsgleichungen* zurückführbar auf die Ausgleichung von *n vermittelnden Beobachtungen mit n — r Unbekannten.*

171. Direkte Lösung des Problems. Außer der eben dargestellten indirekten Lösung läßt das Problem eine direkte Lösung zu, die auf dem Prinzipe beruht, die Summe $[\lambda\lambda]$ so klein zu machen, als es die Bedingungen (4) gestatten. Analytisch kommt dies darauf hinaus, das System der λ so zu bestimmen, daß

$$[\lambda\lambda] \text{ ein Minimum}$$

werde unter gleichzeitiger Erfüllung der Gleichungen:

$$[a\lambda] + w_1 = 0, \quad [b\lambda] + w_2 = 0, \quad \cdots \quad [g\lambda] + w_r = 0; \tag{4}$$

dies aber ist dann geschehen, wenn man die λ und die Größen $k_1, k_2, \cdots k_r$ so bestimmt, daß die Funktion

$$[\lambda\lambda] - 2k_1\{[a\lambda] + w_1\} - 2k_2\{[b\lambda] + w_2\} - \cdots - 2k_r\{[g\lambda] + w_r\}$$

ein absolutes Minimum erlangt und die Gleichungen (4) befriedigt werden.

Die Ausführung dieser Forderung, wobei wir uns auf den Fall dreier Bedingungsgleichungen beschränken, führt auf das Gleichungssystem:

$$\begin{aligned}
\lambda_1 &= a_1k_1 + b_1k_2 + c_1k_3 \\
\lambda_2 &= a_2k_1 + b_2k_2 + c_2k_3 \\
&\cdots\cdots\cdots\cdots \\
\lambda_n &= a_nk_1 + b_nk_2 + c_nk_3
\end{aligned} \tag{6}$$

$$[a\lambda] + w_1 = 0, \quad [b\lambda] + w_2 = 0, \quad [c\lambda] + w_3 = 0, \tag{4*}$$

das aus $n + 3$ Gleichungen besteht und daher zur Lösung der Aufgabe gerade ausreicht.

Die Hilfsgrößen k_1, k_2, $\cdots k_r$ heißen *Korrelaten*,[1]) die Gleichungen (6), welche die λ durch sie ausdrücken, die *Korrelatengleichungen*.

Durch Einsetzung von (6) in (4*) ergeben sich die *Normalgleichungen*:

$$\begin{aligned}
[aa]k_1 + [ab]k_2 + [ac]k_3 + w_1 &= 0 \\
[ab]k_1 + [bb]k_2 + [bc]k_3 + w_2 &= 0 \\
[ac]k_1 + [bc]k_2 + [cc]k_3 + w_3 &= 0,
\end{aligned} \tag{7}$$

die ähnlich gebaut sind wie die Normalgleichungen bei vermittelnden Beobachtungen und zur Berechnung der Korrelaten dienen.

Ist diese vollzogen, so gibt die Eintragung der gefundenen Werte in (6) das vorteilhafteste System der λ, und damit ist auch das vorteilhafteste System der Werte der Unbekannten gefunden:

$$x_1 = l_1 + \lambda_1, \quad x_2 = l_2 + \lambda_2, \quad \cdots \quad x_n = l_n + \lambda_n. \tag{8}$$

Zur Auflösung der Normalgleichungen wird man, sobald ihre Zahl größer und ihre Koeffizienten vielziffrige Zahlen sind, das in Nr. 161 erläuterte Verfahren anwenden, das schließlich zu dem System *reduzierter Normalgleichungen* hinführt:

$$\begin{aligned}
k_1 + \frac{[ab]}{[aa]}k_2 + \frac{[ac]}{[aa]}k_3 + \frac{w_1}{[aa]} &= 0 \\
k_2 + \frac{[bc\cdot 1]}{[bb\cdot 1]}k_3 + \frac{[w_2\cdot 1]}{[bb\cdot 1]} &= 0 \\
k_3 + \frac{[w_3\cdot 2]}{[cc\cdot 2]} &= 0.
\end{aligned} \tag{9}$$

Dabei haben die neu eingeführten Symbole $[w_2 \cdot 1]$, $[w_3 \cdot 1]$, $[w_3 \cdot 2]$ folgende Bedeutung:

$$\begin{aligned}
[w_2 \cdot 1] &= w_2 - \frac{[ab]}{[aa]} w_1 \\
[w_3 \cdot 1] &= w_3 - \frac{[ac]}{[aa]} w_1 \\
[w_3 \cdot 2] &= [w_3 \cdot 1] - \frac{[bc\cdot 1]}{[bb\cdot 1]} [w_2 \cdot 1].
\end{aligned} \tag{10}$$

Es erübrigt noch die *Genauigkeitsbestimmung*. Was den *mittleren Fehler einer Beobachtung* anlangt, so ergibt sich dieser aus der Bemerkung am Schlusse der vorigen Nummer, wonach die Aufgabe der Ausgleichung von n Größen mit r Bedingungsgleichungen äquivalent

---

1) Helmert, Die Ausgleichungsrechnung usw., 2. Aufl., Leipzig 1907, p. 228, bezeichnet die Ausgleichung bedingter Beobachtungen als „Korrelaten-Ausgleichung."

ist der Ausgleichung von n vermittelnden Beobachtungen mit $u = n - r$ Unbekannten; mithin ist nach Nr. 159, (20):

$$\mu = \sqrt{\frac{[\lambda\lambda]}{r}}. \tag{11}$$

Dieser mittlere Fehler bezieht sich aber auf die einzelne Beobachtung als solche, nicht aber auf die *ausgeglichene Beobachtung*, die durch die Ausgleichung mit den andern in Kombination getreten ist. Um z. B. den mittleren Fehler von

$$x_1 = l_1 + \lambda_1$$

zu bestimmen, wäre der folgende Weg einzuschlagen: Nach (6) stellt sich λ_1 durch die Korrelaten dar; für diese ergeben sich aus (7) lineare Ausdrücke der w; die w aber sind nach (2) Funktionen der unabhängig erhaltenen Beobachtungen $l_1, l_2, \cdots l_n$; folglich ist x_1 auch eine Funktion dieser letzteren Größen, und sein mittlerer Fehler ist mit Benützung von μ nach den Vorschriften der Nr. 155 zu berechnen. Wir begnügen uns mit dieser Angabe des wesentlichen Gedankenganges, der immer zu befolgen ist, wenn es sich um die Genauigkeitsbestimmung einer Funktion der ausgeglichenen Beobachtungen handelt, und verweisen bezüglich der systematischen Durchführung auf die spezielle Literatur[1]).

Eine summarische *Rechnungskontrolle* besteht in der zweifachen Bestimmung von $[\lambda\lambda]$. Außer der direkten, welche in dem Quadrieren der einzelnen λ und darauf folgender Summenbildung besteht, gibt es eine indirekte Berechnung jener Summe. Multipliziert man nämlich jede der Gleichungen (6) mit dem entsprechenden λ und bildet hierauf die Summe unter Rücksichtnahme auf (4*), so ergibt sich:

$$[\lambda\lambda] = -[kw]. \tag{12}$$

172. Fortsetzung. Beobachtungen ungleicher Genauigkeit. Sind die Beobachtungen $l_1, l_2, \cdots l_n$ ungleich genau, und ist das Verhältnis ihrer Genauigkeit durch die Gewichte $p_1, p_2, \cdots p_n$ gegeben, so bilden die Forderungen:

$$[p\lambda\lambda] \text{ ein Minimum,}$$

$$[a\lambda] + w_1 = 0,\ [b\lambda] + w_2 = 0,\ [c\lambda] + w_3 = 0 \tag{4**}$$

die Grundlage der Lösung.

Durch Differentiation der Funktion

$$[p\lambda\lambda] - 2k_1\{[a\lambda + w_1\} - 2k_2\{[b\lambda + w_2\} - 2k_3\{[c\lambda] + w_3\}$$

bezüglich der einzelnen λ und Nullsetzen der Ableitungen ergeben sich die *Korrelatengleichungen*:

---

1) Z. B. bei Helmert, Die Ausgleichungsrechnung usw., 2. Aufl., Leipzig 1907, p. 236 ff.

$$\begin{aligned}
\lambda_1 &= \frac{a_1}{p_1} k_1 + \frac{b_1}{p_1} k_2 + \frac{c_1}{p_1} k_3 \\
\lambda_2 &= \frac{a_2}{p_2} k_1 + \frac{b_2}{p_2} k_2 + \frac{c_2}{p_2} k_3 \\
&\cdots\cdots\cdots \\
\lambda_n &= \frac{a_n}{p_n} k_1 + \frac{b_n}{p_n} k_2 + \frac{c_n}{p_n} k_3 ;
\end{aligned} \tag{6*}$$

ihre Einführung in die Bedingungsgleichungen (4**) liefert die ***Normalgleichungen***:

$$\begin{aligned}
\left[\frac{aa}{p}\right] k_1 + \left[\frac{ab}{p}\right] k_2 + \left[\frac{ac}{p}\right] k_3 + w_1 = 0 \\
\left[\frac{ab}{p}\right] k_1 + \left[\frac{bb}{p}\right] k_2 + \left[\frac{bc}{p}\right] k_3 + w_2 = 0 \\
\left[\frac{ac}{p}\right] k_1 + \left[\frac{bc}{p}\right] k_2 + \left[\frac{cc}{p}\right] k_3 + w_3 = 0,
\end{aligned} \tag{7*}$$

deren Auflösung nach demselben Schema vor sich geht wie im Falle gleich genauer Beobachtungen. Aus den Korrelaten berechnen sich nach (6*) die Verbesserungen und mit diesen die ausgeglichenen Beobachtungen:

$$x_1 = l_1 + \lambda_1, \quad x_2 = l_2 + \lambda_2, \cdots \quad x_n = l_n + \lambda_n. \tag{8*}$$

Für den *mittleren Fehler der Gewichtseinheit* gilt nun die aus (7*), Nr. 160 hervorgehende Formel:

$$\mu = \sqrt{\frac{[p\lambda\lambda]}{r}}, \tag{11*}$$

wenn r die Anzahl der Bedingungsgleichungen bedeutet.

Der Beobachtung l_ν an sich, also vor der Ausgleichung, kommt der mittlere Fehler

$$\mu_\nu = \frac{\mu}{\sqrt{p_\nu}}$$

zu; um den mittleren Fehler der ausgeglichenen Beobachtung $x_\nu = l_\nu + \lambda_\nu$ zu bestimmen, müßte der am Schlusse der vorigen Nummer angedeutete Weg eingeschlagen werden.

Die Summe $[p\lambda\lambda]$ kann auch hier zur Kontrolle auf indirektem Wege bestimmt werden; es ergibt sich aus den Gleichungen (6*) mit Zuhilfenahme von (4**) so wie im früheren Falle:

$$[p\lambda\lambda] = -[kw]. \tag{12*}$$

173. Beispiel LXV. *Winkelausgleichung in einem Dreieck.* a) Für die Winkel des Triangulierungsdreiecks MOL sind durch gleich genaue Beobachtungen die Werte:

$$l_1 = 69^0\ 35'\ 12'',61$$
$$l_2 = 67\ \ 46\ \ \ 2\ ,28$$
$$l_3 = 42\ \ 38\ \ 45\ ,97$$

gefunden worden,[1]) deren Summe

$$l_1 + l_2 + l_3 = 180^0\ 0'\ 0'',86$$

beträgt. Da vermöge der Ausdehnung des Dreiecks sein sphärischer Exzeß zu 0'',41 gefunden wurde, ist $180^0\,0'\,0'',41$ die theoretische Winkelsumme, folglich $0,86 - 0,41 = 0'',45$ der Widerspruch, und

$$\lambda_1 + \lambda_2 + \lambda_3 + 0'',45 = 0$$

die von den Verbesserungen zu befriedigende Bedingungsgleichung.

Da $a_1 = a_2 = a_3 = 1$ und alle weiteren Koeffizienten in der allgemeinen Darstellung Null sind, lauten die Korrelatengleichungen:

$$\lambda_1 = \lambda_2 = \lambda_3 = k;$$

die einzige Normalgleichung:

$$3k + 0'',45 = 0$$

gibt

$$k = -0'',15;$$

somit ist auch

$$\lambda_1 = \lambda_2 = \lambda_3 = -0'',15.$$

Der Widerspruch ist also im Falle gleicher Genauigkeit der Beobachtungen auf die drei Winkel gleichmäßig zu verteilen; die ausgeglichenen Beobachtungen:

$$x_1 = 69^0\ 35'\ 12'',46$$
$$x_2 = 67\ \ 46\ \ \ 2\ ,13$$
$$x_3 = 42\ \ 38\ \ 45\ ,82$$

ergeben die durch die Theorie geforderte Summe.

Weil $[\lambda\lambda] = 3 \cdot 0,15^2 = 0,0675$, so ist der mittlere Fehler einer unausgeglichenen Beobachtung

$$\mu = \sqrt{\frac{0,0675}{1}} = 0'',26.$$

Um den mittleren Fehler der ausgeglichenen Beobachtung $x_1 = l_1 + \lambda_1$ zu erhalten, beachte man, daß $\lambda_1 = k$, weiter $k = -\frac{w}{3}$ und

1) Die europäische Längengradmessung. Veröffentlich. d. Königl. Preuß. Geod. Inst., Berlin 1893, p. 260.

$w = l_1 + l_2 + l_3 - S$ ist, wenn S die theoretische Winkelsumme bedeutet; mithin ist

$$x_1 = \frac{S}{3} + \frac{2}{3} l_1 - \frac{1}{3} l_2 - \frac{1}{3} l_3$$

und daher nach (2), Nr. 155:

$$\mu_{x_1} = \mu \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = \mu \sqrt{\frac{2}{3}} = 0'',21;$$

ebenso groß ergeben sich μ_{x_2} und μ_{x_3}.

b) In einem Dreieck wurden die Winkel durch Repetition gemessen; es ergaben sich die Resultate:

$$\begin{aligned} l_1 &= 81^0\ 21'\ 43'',36 \text{ aus } 70 \text{ Repetitionen} \\ l_2 &= 25\ \ 16\ \ 28\ ,85\ \ \text{„}\ \ 101 \quad \text{„} \\ l_3 &= 73\ \ 21\ \ 46\ ,35\ \ \text{„}\ \ \ 85 \quad \text{„} \quad . \end{aligned}$$

Der sphärische Exzeß berechnete sich mit 0'',138; daher zeigt

$$l_1 + l_2 + l_3 = 179^0\ 59'\ 58'',56$$

gegenüber der theoretischen Summe 180° 0' 0'',138 einen Widerspruch $w = -1'',578$, so daß die Bedingungsgleichung

$$\lambda_1 + \lambda_2 + \lambda_3 - 1'',578 = 0$$

zu befriedigen sein wird.

Die Repetitionszahlen bilden zugleich die Gewichte der Beobachtungen, und da in der einzigen Bedingungsgleichung alle Koeffizienten gleich sind, so lauten die Korrelatengleichungen:

$$\lambda_1 = \frac{k}{p_1}, \quad \lambda_2 = \frac{k}{p_2}, \quad \lambda_3 = \frac{k}{p_3}$$

und die Normalgleichung:

$$\left[\frac{1}{p}\right] k + w = 0,$$

so daß

$$k = -\frac{w}{\left[\frac{1}{p}\right]} = \frac{1'',578}{0,03595} = 43'',894.$$

Mithin ist

$$\lambda_1 = \frac{43'',894}{70} = 0'',627, \quad \lambda_2 = \frac{43'',894}{101} = 0'',435, \quad \lambda_3 = \frac{43'',894}{85} = 0'',516,$$

und die ausgeglichenen Winkel

$$\begin{aligned} x_1 &= l_1 + \lambda_1 = 81^0 21' 43'',987 \\ x_2 &= l_2 + \lambda_2 = 25\ \ 16\ \ 29\ ,285 \\ x_3 &= l_3 + \lambda_3 = 73\ \ 21\ \ 46\ ,866 \end{aligned}$$

ergeben die theoretische Summe 180° 0' 0'',138.

Durch direkte Ausrechnung findet man

$$[p\lambda\lambda] = 69{,}2648,$$

mittels der Kontrollgleichung

$$[p\lambda\lambda] = 43{,}894 \cdot 1{,}578 = 69{,}2647;$$

daraus bestimmt sich der mittlere Fehler der Gewichtseinheit:

$$\mu = \sqrt{69{,}2648} = 8'',32;$$

die unausgeglichenen Beobachtungen haben somit die mittleren Fehler:

$$\mu_1 = \frac{\mu}{\sqrt{p_1}} = 1'',00, \quad \mu_2 = \frac{\mu}{\sqrt{p_2}} = 0'',83, \quad \mu_3 = \frac{\mu}{\sqrt{p_3}} = 0'',90.$$

Um den mittleren Fehler des ausgeglichenen

$$x_1 = l_1 + \lambda_1$$

zu finden, beachte man, daß $\lambda_1 = \frac{k}{p_1}$, weiter $k = -\frac{w}{\left[\frac{1}{p}\right]}$, und $w = l_1 + l_2 + l_3 - S$, wenn S die theoretische Winkelsumme bezeichnet; mithin ist

$$x_1 = \frac{S}{p_1\left[\frac{1}{p}\right]} + \left(1 - \frac{1}{p_1\left[\frac{1}{p}\right]}\right) l_1 - \frac{1}{p_1\left[\frac{1}{p}\right]} l_2 - \frac{1}{p_1\left[\frac{1}{p}\right]} l_3,$$

und nach (2), Nr. 155:

$$\mu_{x_1} = \sqrt{\left(1 - \frac{1}{p_1\left[\frac{1}{p}\right]}\right)^2 \mu_1^2 + \frac{1}{p_1^2\left[\frac{1}{p}\right]^2}\mu_2^2 + \frac{1}{p_1^2\left[\frac{1}{p}\right]^2}\mu_3^2}$$

$$= \mu \sqrt{\frac{\left(\frac{1}{p_2} + \frac{1}{p_3}\right)\frac{1}{p_1}}{\left[\frac{1}{p}\right]}} = 0'',77;$$

ebenso ergibt sich

$$\mu_{x_2} = \mu \sqrt{\frac{\left(\frac{1}{p_1} + \frac{1}{p_3}\right)\frac{1}{p_2}}{\left[\frac{1}{p}\right]}} = 0'',70, \qquad \mu_{x_3} = \mu \sqrt{\frac{\left(\frac{1}{p_1} + \frac{1}{p_2}\right)\frac{1}{p_3}}{\left[\frac{1}{p}\right]}} = 0'',73.$$

Wie in beiden Fällen zu bemerken, werden die mittleren Fehler der Winkel durch die Ausgleichung vermindert.

174. Beispiel LXVI. *Ausgleichung eines Vierecks.* In dem vierpunktigen Dreiecksnetz $ABCD$, Fig. 18, sind für die durch Bogen und Nummern bezeichneten Winkel folgende gleich genaue Werte erhalten worden:

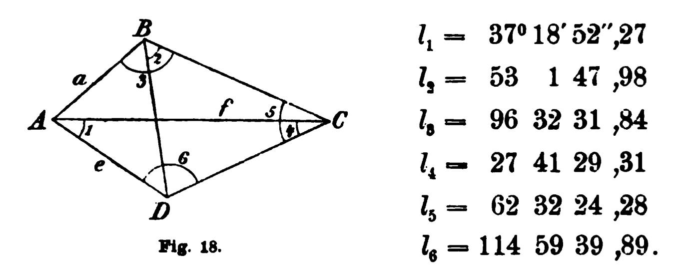


Fig. 18.

$l_1 = 37^0\,18'\,52'',27$
$l_2 = 53\ \ 1\ 47\ ,98$
$l_3 = 96\ 32\ 31\ ,84$
$l_4 = 27\ 41\ 29\ ,31$
$l_5 = 62\ 32\ 24\ ,28$
$l_6 = 114\ 59\ 39\ ,89.$

Da zur Festlegung der relativen Lage von vier Punkten vier Winkel ausreichen, so sind zwei Messungen überschüssig; daher werden zwischen den Winkeln zwei unabhängige Bedingungsgleichungen sich aufstellen lassen.

Die eine ergibt sich aus dem Dreieck ACD, weil darin alle drei Winkel gemessen sind; sie lautet:

$$\lambda_1 + \lambda_4 + \lambda_6 + w_1 = 0 \tag{1}$$

mit

$$w_1 = l_1 + l_4 + l_6 - (180^0 + E)$$

und heißt eine *Winkelgleichung.*

Würde man nach Ausgleichung dieser drei Winkel das Dreieck ACD konstruieren, durch C einen Strahl unter dem Winkel l_5 gegen CD ziehen, in diesem Strahl den Punkt B so bestimmen, daß der Winkel CBD gleich l_2 wird, endlich durch B einen Strahl unter dem Winkel l_3 gegen BC führen, so würde dieser wegen der Fehlerhaftigkeit der Winkel l_5, l_2, l_3 nicht durch A gehen, das Viereck würde sich nicht schließen. Die Winkel müssen daher noch derart abgeändert werden, daß für eine Länge, die sich mittels der gemessenen Winkel auf zwei verschiedenen Wegen rechnen läßt, beidemal derselbe Wert erhalten werde.

Für f ergibt sich aus dem Dreieck ACD der Ausdruck:

$$f = e\,\frac{\sin(l_6 + \lambda_6)}{\sin(l_4 + \lambda_4)};$$

aus dem Dreieck ABC findet man zunächst:

$$f = a\,\frac{\sin(l_3 + \lambda_3)}{\sin(l_5 + \lambda_5 - l_4 - \lambda_4)},$$

und für a aus dem Dreieck ABD:

$$a = e\,\frac{\sin(l_2 + \lambda_2 + l_5 + \lambda_5 + l_6 + \lambda_6 - 180^0)}{\sin(l_3 + \lambda_3 - l_2 - \lambda_2)};$$

daher ist auf dem andern Wege

$$f = e\,\frac{\sin(l_3 + \lambda_3)\sin(l_2 + \lambda_2 + l_5 + \lambda_5 + l_6 + \lambda_6 - 180^0)}{\sin(l_5 + \lambda_5 - l_4 - \lambda_4)\sin(l_3 + \lambda_3 - l_2 - \lambda_2)}.$$

Durch Gleichsetzung beider Resultate ergibt sich die zweite Bedingungsgleichung, eine sogenannte *Seitengleichung:*

$$\frac{\sin(l_5+\lambda_5-l_4-\lambda_4)\sin(l_6+\lambda_6)\sin(l_3+\lambda_3-l_2-\lambda_2)}{\sin(l_3+\lambda_3)\sin(l_4+\lambda_4)\sin(l_2+\lambda_2+l_5+\lambda_5+l_6+\lambda_6-180^0)}=1. \qquad (2)$$

Die Gleichung (1) besitzt bereits die lineare Form

$$[a\lambda]+w_1=0,$$

und zwar ist (der Exzeß des Dreiecks = 3'',74)

$$a_1=a_4=a_6=1,\quad a_2=a_3=a_5=0,$$

$$w_1=180^0\,0'\,1'',47-180^0\,0'\,3'',74=-2'',27.$$

Bei (2) muß diese Form erst hergestellt werden durch Entwicklung nach den λ. Logarithmiert man zu diesem Zwecke und setzt:

$$w_2=\log\sin(l_5-l_4)+\log\sin l_6+\log\sin(l_3-l_2)$$

$$-\log\sin l_3-\log\sin l_4-\log\sin(l_2+l_5+l_6-180^0),$$

so sind die Koeffizienten der zweiten Bedingungsgleichung

$$[b\lambda]+w_2=0$$

die Ableitungen von w_2 nach den l, also z. B.

$$b_1=\frac{\partial w_2}{\partial l_1}=0$$

$$b_2=\frac{\partial w_2}{\partial l_2}=-M(\operatorname{cotg}(l_3-l_2)+\operatorname{cotg}(l_2+l_5+l_6-180^0))$$

usw.; M bedeutet den Modul des gemeinen Logarithmensystems. Diese umständliche Berechnung der Koeffizienten kann mit Hilfe der Differenzen in der logarithmisch-trigonometrischen Tafel umgangen werden; ist nämlich l ein endlicher und λ ein sehr kleiner, in Sekunden ausgedrückter Winkel, so kann mit zureichender Schärfe

$$\log\sin(l+\lambda)=\log\sin l+\Delta\cdot\lambda$$

gesetzt werden, wenn Δ die Differenz zu $\log\sin l$ pro 1'' bedeutet. Demnach hat man zur Berechnung von w_2 selbst *und* der Koeffizienten folgenden Ansatz:

$$\begin{array}{lrl}
\log\sin(l_5-l_4+\lambda_5-\lambda_4) & = & 9{,}7569478.1 \qquad -30{,}2\,\lambda_4+30{,}2\,\lambda_5 \\
\log\sin(l_6+\lambda_6) & = & 9{,}9572954.1 \qquad -9{,}9\,\lambda_6 \\
\log\sin(l_3+l_2+\lambda_3+\lambda_2) & = & \underline{9{,}8379094.7}-22{,}2\,\lambda_2+22{,}2\,\lambda_3 \\
 & & 29{,}5521526.9 \\
\log\sin(l_3+\lambda_3) & = & 9{,}9971627.6 \qquad -2{,}4\,\lambda_3 \\
\log\sin(l_4+\lambda_4) & = & 9{,}6671823.3 \qquad +40{,}2\,\lambda_4 \\
\log\sin\begin{pmatrix} l_2+l_5+l_6-180^0 \\ +\lambda_2+\lambda_5+\lambda_6 \end{pmatrix} & = & \underline{9{,}8878085.2}+17{,}3\,\lambda_2 \qquad +17{,}3\,\lambda_5+17{,}3\,\lambda_6 \\
 & & 29{,}5521536.1
\end{array}$$

Durch Subtraktion der unteren Gruppe von der oberen ergibt sich die Bedingungsgleichung:

$$-39{,}5\lambda_2 + 24{,}6\lambda_3 - 70{,}4\lambda_4 + 12{,}9\lambda_5 - 27{,}2\lambda_6 - 9{,}2 = 0.$$

Dividiert man hier, um auf kleinere Zahlen zu kommen, durch 10, so kann man die *Bedingungsgleichungen* endgültig schreiben:

$$\lambda_1 \qquad\qquad + \lambda_4 \qquad\qquad + \lambda_6 - 2{,}27 = 0$$
$$-3{,}95\lambda_2 + 2{,}46\lambda_3 - 7{,}04\lambda_4 + 1{,}29\lambda_5 - 2{,}72\lambda_6 - 0{,}92 = 0.$$

Dies führt zu den *Korrelatengleichungen:*

$$\begin{aligned} \lambda_1 &= k_1 \\ \lambda_2 &= \quad -3{,}95k_2 \\ \lambda_3 &= \qquad 2{,}46k_2 \\ \lambda_4 &= k_1 - 7{,}04k_2 \\ \lambda_5 &= \qquad 1{,}29k_2 \\ \lambda_6 &= k_1 - 2{,}72k_2: \end{aligned}$$

daraus ergeben sich die *Normalgleichungen:*

$$3{,}00k_1 - 9{,}76k_2 - 2{,}27 = 0$$
$$-9{,}76k_1 + 80{,}28k_2 - 0{,}92 = 0;$$

ihre Auflösung liefert:

$$k_1 = 1{,}3139$$
$$k_2 = 0{,}1712;$$

hiermit berechnen sich aus den Korrelatengleichungen:

$$\begin{aligned} \lambda_1 &= +1'',314 \\ \lambda_2 &= -0\ ,677 \\ \lambda_3 &= +0\ ,420 \\ \lambda_4 &= +0\ ,108 \\ \lambda_5 &= +0\ ,220 \\ \lambda_6 &= +0\ ,848. \end{aligned}$$

Die *ausgeglichenen Winkel* des Netzes sind sonach:

$$\begin{aligned} x_1 &= l_1 + \lambda_1 = \ \ 37^0 18' 53'',584 \\ x_2 &= l_2 + \lambda_2 = \ \ 53\ \ \ 1\ 47\ ,303 \\ x_3 &= l_3 + \lambda_3 = \ \ 96\ 32\ 32\ ,260 \\ x_4 &= l_4 + \lambda_4 = \ \ 27\ 41\ 29\ ,418 \\ x_5 &= l_5 + \lambda_5 = \ \ 62\ 32\ 24\ ,500 \\ x_6 &= l_6 + \lambda_6 = 114\ 59\ 40\ ,738; \end{aligned}$$

sie führen zu einer widerspruchsfreien Berechnung desselben.

Auf direktem Wege findet man

$$[\lambda\lambda] = 3{,}1405,$$

während sich indirekt, durch Benützung der Korrelaten und der Widersprüche, in genügender Übereinstimmung

$$[\lambda\lambda] = 1{,}3139 \cdot 2{,}27 + 0{,}1712 \cdot 0{,}92 = 3{,}1400$$

ergibt; hiernach ist der mittlere Fehler einer Winkelmessung:

$$\mu = \sqrt{\frac{3{,}14}{2}} = 1'',25.$$

An ihn hat die Genauigkeitsbestimmung der ausgeglichenen Winkel anzuschließen.[1])

---

1) Bezüglich der Anwendungen der Methode der kleinsten Quadrate, insbesondere auf Probleme der Geodäsie, aber auch Fragen physikalischer Natur, auf Interpolation u. a. sei auf die nach dieser Richtung sehr reichhaltige 2. Auflage von Helmert, Die Ausgleichungsrechnung etc., Leipzig 1907, verwiesen.

Dritter Teil.

Kollektivmaßlehre.

§ 1. Kollektivgegenstände und die Mittel zu ihrer Beschreibung.

175. Begriffsbestimmung eines Kollektivgegenstandes. Naturobjekten, Naturerscheinungen, Erzeugnissen menschlicher Tätigkeit, physischer wie geistiger, sozialen Vorgängen gegenüber erwachsen der Wissenschaft mannigfache Aufgaben.

Soll von einem Ding, einer Erscheinung, einem Vorgang, kurz von einem einzelnen Wahrnehmungsobjekt eine Beschreibung gegeben werden, so handelt es sich um die Angabe solcher Merkmale, auf Grund deren das Objekt erkannt und von andern gleichartigen unterschieden werden kann; es ist eine *Individualbeschreibung*, die auf solche Weise zustande kommt.

Um eine Art oder Gattung von Objekten zu beschreiben, hat man einen solchen Komplex von Merkmalen zusammenzustellen, mit dessen Hilfe es möglich ist, über die Zugehörigkeit eines Objekts zur Gattung zu entscheiden; eine solche Beschreibung, eine *Gattungsbeschreibung*, erfordert die vergleichende Betrachtung einer Vielheit von Objekten; sie ist das Produkt einer Abstraktion, weil von manchen Eigenschaften, die bei der Individualbeschreibung notwendig wären, abgesehen wird. Die Gattungsbeschreibung bestimmt den Begriff der Gattung.

Objekte einer Gattung lassen häufig eine Klassifizierung nach einzelnen Merkmalen oder Merkmalkomplexen zu.

Eine Menge von Objekten werde als gleichartig angesehen, oder sie sei es wirklich, in allen Merkmalen bis auf eines, das von Objekt zu Objekt variiert; ist dieses Merkmal *zahlenmäßig* darstellbar, so eignet sich die Menge zum Gegenstande einer mathematischen Bearbeitung, deren erster Schritt darin bestehen möge, daß man die Objekte der Menge nach diesem Merkmal *ordnet*. Dabei wird, im Gegensatze zur Gattungs- und Individualbeschreibung, von allen übrigen Merkmalen abstrahiert und nur das eine zunächst individuell behandelt. Würde auch noch von dem einen Merkmal abgesehen, so hätte man es mit einer Menge gleichberechtigter Einheiten zu tun, wie sie dem *Zählprozeß* zugrunde liegen.

Eine Menge von gleichartigen Objekten, die in bezug auf ein veränderliches, zahlenmäßig darstellbares Merkmal geordnet werden können, bezeichnet man als einen Kollektivgegenstand.

Der so festgesetzte Begriff ist sehr umfassend, darum die große Tragweite der auf ihm aufgebauten *Kollektivmaßlehre.* Die nachfolgende kleine Zusammenstellung soll die Vielseitigkeit des Begriffs wenigstens andeuten. Vorher mögen jedoch einige seiner Detailbestimmungen erörtert werden.

Die einzelnen Objekte nennt man *Exemplare* des Kollektivgegenstandes oder *Glieder* der Kollektivreihe, ihre Anzahl (m) den *Umfang* der Reihe. Die Bildung der Glieder ist entweder durch die Natur der Sache gegeben oder kann auch willkürlich geschehen; die Zusammenstellung führt Beispiele für beides an.

Kollektivgegenstand	Exemplar	Ordnendes Merkmal
1. Ziehungen aus einer Urne mit weißen und schwarzen Kugeln	Serie von s Ziehungen	Häufigkeit weißer Kugeln
2. Komposite Blüten einer Art	Einzelblüte	Zahl der Staubfäden
3. Prosasprache eines bestimmten Idioms	Druckseite eines darin geschriebenen Buches	Häufigkeit des Auftretens eines bestimmten Buchstabens
4. Geburten bestimmter Herkunft und aus bestimmter Zeitperiode	Serie von s Geburtsfällen	Häufigkeit männlicher Geburten
5. Messungen einer physischen Größe	Ergebnis der Einzelmessung	Messungsfehler
6. Temperaturbeobachtungen an einem Orte	Bestimmter Tag eines Jahres	Tagesmittel (Mittel zwischen höchster und niederster Temperatur)
7. Personen männlichen Geschlechts und zwischen bestimmten Altersgrenzen etc.	Individuum	Körperlänge; Brustumfang; Schädelumfang; Kopflänge; Kopfbreite; Körpergewicht etc.
8. Gestorbene, männlichen Geschlechts, bestimmter Herkunft etc.	Individuum	Sterbealter
9. Personen, bestimmten Geschlechts, bestimmter Nationalität etc.	Individuum	Farbe der Augen; Haarfarbe o. dgl.

Die Voraussetzung einer zahlenmäßigen Darstellung ist unmittelbar erfüllt, wenn das ordnende Merkmal quantitativer Natur ist, wenn es also durch Zählung, Messung, Wägung u. dgl. festgestellt werden kann. Aber auch bei qualitativen Merkmalen läßt sich eine zahlenmäßige Darstellung künstlich erzielen; fertigt man nämlich eine Skala der Abarten, Nuancen, Stufen — oder wie die Disjunktion im ein-

zelnen Falle heißen möge — der ordnenden Eigenschaft an und belegt sie mit Zahlen, so ist damit jedem Exemplar des Kollektivgegenstandes eine Zahl zugeordnet, und darin liegt die Vorbedingung der Ordnungsmöglichkeit.

Die in der Definition aufgestellte Forderung der Gleichartigkeit der Objekte kann in verschiedenem Maße erfüllt sein. Je weiter man sie treibt, um so schwieriger wird es sein, eine Kollektivreihe von genügend großem Umfang zu bilden; und auf den Umfang kommt es wesentlich an, weil die Ergebnisse der später vorzunehmenden Untersuchungen desto wertvoller sind, je mehr Grade, Stufen o. dgl. des ordnenden Merkmals in der Reihe vertreten sind. Ein zu geringer Grad von Gleichartigkeit kann aber ebenso wie ein zu kleiner Umfang das Ergebnis wertlos machen. Faßt man z. B., wenn die Körperlänge das ordnende Merkmal ist, Kinder und Erwachsene, männliche und weibliche Personen zusammen, so darf das Resultat ernstliches Interesse nicht beanspruchen; hingegen kann ihm auch praktische Bedeutung zukommen, wenn die Personen nach bestimmten Grundsätzen ausgewählt sind (z. B. Rekruten eines großen Aushebebezirks).

176. Argument. Stetige und unstetige Kollektivgegenstände. Die veränderliche Zahl x, die in ihren individuellen Werten das ordnende Merkmal kennzeichnet, nennt man das *Argument* des Kollektivgegenstandes.

Es kann in der Natur des letzteren liegen, daß x nur einer Reihe diskreter Werte fähig ist. Dies tritt immer ein, wenn das Merkmal durch Zählung festgestellt wird, wie in den Beispielen 2, 3. Wählt man in dem Beispiel 1 die *absolute* Häufigkeit der weißen Kugeln als Argument, so ist x aller Werte aus der Reihe $0, 1, 2, \cdots s$ fähig; dagegen aller Werte aus der Reihe $0, \frac{1}{s}, \frac{2}{s}, \cdots 1$, wenn die *relative* Häufigkeit zum Argument genommen wird; ähnliches gilt im Beispiel 4.

In allen diesen Fällen ist das Argument eine *unstetige* Variable, und diese Bezeichnung überträgt man auch auf den Kollektivgegenstand selbst.

Man spricht hingegen von einem *stetigen* Kollektivgegenstand, wenn das Argument seiner Natur nach als stetige Variable aufzufassen ist, indem es eine extensive Größe ausdrückt; Beleg hierfür sind die Beispiele 5—8.

Das Ordnen eines unstetigen Kollektivgegenstandes besteht darin, daß man jedem möglichen *Einzelwert* x_i des Arguments die Zahl z_i zuordnet, welche die Anzahl der Exemplare dieses Einzelwerts angibt.

Um einen stetigen Kollektivgegenstand zu ordnen, hat man das

G. F. Lipps, Theorie der Kollektivgegenstände, Leipzig 1902, p. 47, bezeichnet die unter einen Argumentwert, beziehungsweise in ein Intervall fallenden Exemplare als *Variante* des Kollektivgegenstandes.

Gebiet der möglichen Argumentwerte in Intervalle (X_i, X_{i+1}) zu teilen und jedem solchen *Wertintervall* die Zahl z_i der ihm angehörenden Exemplare zuzuordnen.[1]) Die natürlichste und für die weitere Behandlung zweckmäßigste Teilung ist die in *gleiche* Intervalle; dann genügt zur Kennzeichnung eines solchen Intervalls ein einziger ihm angehörender Argumentwert, und als solcher wird die Intervall*mitte* genommen; man ordnet dann die Zahl z_i statt dem Intervall seiner Mitte $\frac{X_i + X_{i+1}}{2} = x_i$ zu. Die Argumentwerte X_i, die die Intervalle voneinander scheiden, heißen *Wechselpunkte*.

Die Zahl z_i drückt die *absolute* Häufigkeit der Exemplare vom Argumentwert x_i in dem einen, in dem durch x_i gekennzeichneten Intervall im andern Fall; man sagt aber Kürze halber, z_i sei die absolute Häufigkeit von x_i. Den Quotienten $\frac{z_i}{m} = y_i$ nennt man die *relative* Häufigkeit.

In bezug auf alle Argumentwerte oder Argumentintervalle gilt:

$$\sum z_i = m, \qquad \sum y_i = 1.$$

Bei unstetigen Kollektivreihen ist das Gebiet der möglichen Argumentwerte meist in bestimmter Weise begrenzt; so reicht, wenn man im Beispiel 4 Serien von je 100 Geburten bildet und die absolute Zahl der männlichen Geburten als Argument wählt, sein Gebiet von 0 bis 100, und von 0 bis 1, wenn man die relative Menge männlicher Geburten (100 als Einheit genommen) als Argument benützt.

Anders liegt die Sache bei stetigen Kollektivreihen; hier ist in der Regel eine bestimmte Begrenzung des Argumentbereichs nicht möglich; in dem Beispiel 7 lassen sich erreichbare und nicht überschreitbare Grenzen der Körperlänge von vornherein nicht angeben; die *beobachteten* Grenzen werden im allgemeinen mit wachsendem m weiter hinausrücken.

Für die Darstellung und zumal für die allgemeine mathematische Behandlung ist es vorteilhaft, die Grenzen in beiden Fällen ins Unendliche rücken zu lassen, indem man festsetzt, daß *außermöglichen* Argumentwerten, bzw. Argumentintervallen, $z = 0$ oder $y = 0$ zuzuordnen ist.

177. Verteilungstafeln. Das Erheben eines Kollektivgegenstandes besteht darin, daß man den jedem Exemplar zukommenden Argumentwert festsetzt und diese Werte, so wie sie sich ergeben, in eine Liste einträgt.

Bei stetigen Kollektivgegenständen wird in der Regel nicht der individuelle, sondern ein *abgerundeter* Argumentwert in die Liste eingetragen, so daß mit der Erhebung die Intervallbildung einhergeht; die Intervallgröße hängt dann mit der Stärke der Abrundung zu-

sammen. Ist beispielsweise bei einer statistischen Erhebung das Alter der Personen ordnendes Argument, so kann die Abrundung ein halbes Jahr auf und ab betragen, so daß unter dem ganzzahligen Alter x_i Personen zusammengefaßt werden, deren Alter zwischen $X_i = x_i - \frac{1}{2}$ und $X_{i+1} = x_i + \frac{1}{2}$ Jahre beträgt. Wird bei Erhebung der Körperlänge von Rekruten auf ganze Zentimeter abgerundet, so hat dies den Sinn, daß mit x_i cm alle Rekruten verzeichnet werden, deren Körperlänge zwischen $X_i = x_i - \frac{1}{2}$ und $X_{i+1} = x_i + \frac{1}{2}$ cm beträgt.

Das Ergebnis der Erhebung ist die *Urliste* des Kollektivgegenstandes. Sie bildet die Grundlage für das Ordnen desselben. Die Zusammenstellung der Argumentwerte bzw. Argumentintervalle in ihrer natürlichen Reihenfolge mit dem niedrigsten beginnend, und der dazugehörigen absoluten Häufigkeitszahlen z oder der relativen y liefert die *Verteilungstafel* des Kollektivgegenstandes als dessen erste Beschreibung.

Die schematische Darstellung der Verteilungstafel eines unstetigen Kollektivgegenstandes ist durch die nachstehende Skizze gegeben:

$$\begin{array}{ccccccc} x_1 & x_2 & x_3 & & x_i & & x_n \\ z_1 & z_2 & z_3 \cdots & & \cdots z_i \cdots & & \cdots z_n \\ (y_1) & (y_2) & (y_3) \cdots & & \cdots (y_i) \cdots & & \cdots (y_n) \end{array}$$

die eines stetigen durch die folgende:

$$\begin{array}{cccccccccccc} X_1 & x_1 & X_2 & x_2 & X_3 & x_3 & X_4 & X_i & x_i & X_{i+1} & X_n & x_n & X_{n+1} \\ & z_1 & & z_2 & & z_3 \cdots & & \cdots & z_i \cdots & & \cdots & z_n & \\ & (y_1) & & (y_2) & & (y_3) \cdots & & \cdots & (y_i) \cdots & & \cdots & (y_n) & \end{array}$$

Errichtet man in den Punkten x_i Lote und trägt darauf nach einem gewählten Maßstab die zugehörigen z_i oder y_i auf, so erhält man ein anschauliches Bild der Verteilung der Argumentwerte auf die Exemplare des Kollektivgegenstandes. Im Falle der Stetigkeit wird das Bild noch vollständiger und anschaulicher, wenn man die Endpunkte der Lote durch einen Linienzug verbindet und so eine *Häufigkeits-* oder *Verteilungskurve* des Kollektivgegenstandes konstruiert. Eine solche Kurve kann auch im Falle der Unstetigkeit als Behelf dienen.

Die *primäre* Verteilungstafel, wie sie sich unmittelbar aus der Urliste ergibt, kann vermöge unzureichenden Umfangs m oder wegen Bildung zu kleiner Intervalle große Unregelmäßigkeiten aufweisen; sie kann auch wegen zu großer Ausdehnung die Übersicht erschweren, für den Zweck, den man verfolgt, zu detailliert sein. Man stellt in solchen Fällen aus der primären eine *reduzierte* Verteilungstafel her,

indem man je zwei, drei oder mehr aufeinanderfolgende Intervalle zu einem neuen, größeren vereinigt; als seinen Vertreter wählt man wieder den Mittelpunkt. Behält man beispielsweise die Wechselpunkte $X_1, X_3, X_5, \cdots$ bei, so kommen die Vertreter $x_1, x_2, \cdots$ der neuen Intervalle an die Stellen $X_2, X_4, \cdots$ und es werden ihnen die Zahlen $z_1 + z_2,\ z_3 + z_4, \cdots$ oder $y_1 + y_2,\ y_3 + y_4, \cdots$ zugeordnet. Man kann aber bei derselben Stärke (Stufe) der Reduktion auch $X_0, X_2, X_4, \cdots$ als Wechselpunkte behalten (X_0 ist dem X_1 in der entsprechenden Entfernung vorgesetzt) und die Vertreter der Intervalle an die Stellen $X_1, X_3, \cdots$ mit den Zahlen $0 + z_1,\ z_2 + z_3, \cdots$ setzen. Die beiden so entstandenen Tafeln unterscheiden sich dann in der *Reduktionslage*. Bei Zusammenfassung von je drei ursprünglichen Intervallen gibt es drei verschiedene Reduktionslagen usw.

Primäre Verteilungstafel.

x_i	Intervall X_i bis X_{i+1}	z_i
18	$17\frac{1}{2} - 18\frac{1}{2}$	1
19	$18\frac{1}{2} - 19\frac{1}{2}$	3
20	$19\frac{1}{2} - 20\frac{1}{2}$	5.5 [1]
21	$20\frac{1}{2} - 21\frac{1}{2}$	9
22	$21\frac{1}{2} - 22\frac{1}{2}$	17.5
23	$22\frac{1}{2} - 23\frac{1}{2}$	31
24	$23\frac{1}{2} - 24\frac{1}{2}$	55
25	$24\frac{1}{2} - 25\frac{1}{2}$	82.5
26	$25\frac{1}{2} - 26\frac{1}{2}$	121
27	$26\frac{1}{2} - 27\frac{1}{2}$	176
28	$27\frac{1}{2} - 28\frac{1}{2}$	232.5
29	$28\frac{1}{2} - 29\frac{1}{2}$	287
30	$29\frac{1}{2} - 30\frac{1}{2}$	336.5
31	$30\frac{1}{2} - 31\frac{1}{2}$	425
32	$31\frac{1}{2} - 32\frac{1}{2}$	502
33	$32\frac{1}{2} - 33\frac{1}{2}$	576.5
34	$33\frac{1}{2} - 34\frac{1}{2}$	671.5
35	$34\frac{1}{2} - 35\frac{1}{2}$	735.5
.	. .	.
.	. .	.
.	. .	.

Reduzierte Verteilungstafel.

x_i	Intervall X_i bis X_{i+1}	z_i
18	$15\frac{1}{2} - 20\frac{1}{2}$	9.5
23	$20\frac{1}{2} - 25\frac{1}{2}$	195
28	$25\frac{1}{2} - 30\frac{1}{2}$	1153
33	$30\frac{1}{2} - 35\frac{1}{2}$	2910.5
38	$35\frac{1}{2} - 40\frac{1}{2}$	4717
43	$40\frac{1}{2} - 45\frac{1}{2}$	6387.5
48	$45\frac{1}{2} - 50\frac{1}{2}$	7555
53	$50\frac{1}{2} - 55\frac{1}{2}$	8133.5
.	. .	.
.	. .	.
.	. .	.

Zur Illustration des Vorgebrachten ist vorstehend ein Bruchstück einer primären Verteilungstafel einer Personengemeinschaft nach dem Sterbealter mitgeteilt und daneben eine reduzierte Verteilungstafel gestellt, die je fünf Jahresintervalle zu einem Altersquinquennium zusammenfaßt; die Reduktionslage ist so gewählt, daß die Mitten der neuen Intervalle auf die Alter 18, 23, $\cdots$ fallen; zu diesem Zwecke

1) Die Bruchteile stammen daher, daß die Zahlen dieser Kolonne das Ergebnis einer rechnerischen Vorbehandlung sind.

müssen dem ersten Intervall $17\frac{1}{2}-18\frac{1}{2}$ der primären Tafel die *leeren* Intervalle $15\frac{1}{2}-16\frac{1}{2}$, $16\frac{1}{2}-17\frac{1}{2}$ vorangestellt werden; ebenso wird am Schlusse, wenn die Teilung nicht aufgeht, die Hinzufügung leerer Intervalle erforderlich sein.

178. Verteilungsfunktion. Mit der Verteilungstafel und Verteilungskurve ist die wissenschaftliche Erforschung der Kollektivgegenstände nicht erschöpft. Eine solche beginnt vielmehr erst dann, wenn man sich die Frage vorlegt, ob die Verteilung verschiedener Kollektivreihen nicht etwa eine einheitliche mathematische Fassung gestatte, durch die es möglich würde, die Verteilung eines bestimmten Kollektivgegenstandes durch gewisse Parameter, Elemente, zu charakterisieren und verschiedene Objektreihen auf ihre Verteilung zu vergleichen.

Die Wahrscheinlichkeitsrechnung führt in den Ergebnissen wiederholter Versuchsreihen (Nr. 86) zu theoretisch konstruierten Kollektivgegenständen, für welche sich die Verteilung a priori angeben läßt, mit der dann die faktische Verteilung verglichen werden kann, und diese Verteilung läßt sich *angenähert* durch eine Funktion beschreiben, die nur von *einem* Parameter (dem Präzisionsmaß) abhängt.

Die Fehlertheorie operiert gleichfalls mit Kollektivreihen, als deren einfachster Vertreter eine Reihe wiederholter Messungen einer und derselben Größe anzusehen ist, und sie bedient sich einer Funktion, die die Verteilung der Fehler nach ihrer Größe beschreibt und gleichfalls nur von einem Parameter abhängt. Auch diese Funktion ist auf apriorischem Wege aus verschiedenen Gesichtspunkten begründet worden und fand vielfache Bestätigung durch die Erfahrung.

An diese Fälle knüpfen denn auch die ersten Ansätze zu einer Kollektivmaßlehre an, die von A. Quetelet stammen[1]; dieser nahm ohne weiteres volle Analogie zwischen der Verteilung von Körperlängen, Brustumfängen u. dgl. einer großen Anzahl erwachsener Personen und der Wiederholungszahlen weißer (oder schwarzer) Kugeln in einer großen Anzahl gleich ausgedehnter Ziehungsreihen aus einer Urne an, die weiße und schwarze Kugeln in gleicher Menge enthält; die Grenzform dieser Verteilung ist das Gaußsche Exponentialgesetz $\varphi(x)$, Nr. 136, Gleichung (6).

Auch der eigentliche Begründer der Kollektivmaßlehre, G. Th. Fechner[2], schloß sich an dieses Gesetz an. Er nahm in seine Definition des Kollektivgegenstandes ausdrücklich das Moment der Zufälligkeit auf, indem er einen solchen als eine Gesamtheit von unbestimmt vielen *nach Zufall* variierenden Exemplaren erklärte, die durch einen Art- oder Gattungsbegriff zusammengehalten werden.

---

1) Lettres sur la théorie des probabilités etc. Bruxelles, 1846.

2) Kollektivmaßlehre. Herausgegeben von G. F. Lipps. Leipzig 1897.

Seine vielfachen Untersuchungen an bereits vorhandenen und an eigens geschaffenen Kollektivgegenständen führten ihn aber weiter. Während nämlich die zu $\varphi(x)$ gehörige Verteilungskurve in bezug auf ihre größte Ordinate Symmetrie aufweist, gelangte Fechner zu Kollektivreihen, die eine von der Symmetrie so stark abweichende Verteilung aufwiesen, daß die Asymmetrie als wesentlich, nicht als eine bloße Störung einer in der Natur der Sache gelegenen Symmetrie erkannt werden mußte. Er glaubte, dieser Tatsache mathematisch dadurch Rechnung tragen zu können, daß er das gewöhnliche (einfache) Gaußsche Gesetz zu einem *zweiseitigen* erweiterte, indem er die Verteilungskurve aus zwei Zweigen zusammensetzte, deren jeder einem Gesetz der Form $\varphi(x)$, aber von ungleichem Präzisionsmaß, folgt und deren Gipfelpunkte zusammenfallen. Nicht in dieser anfechtbaren Erweiterung liegt Fechners Verdienst, sondern in der Schaffung einer Systematik für die Untersuchung von Kollektivgegenständen, in der Einführung einer Terminologie für diesen Zweig der angewandten Mathematik. Wiewohl er ihn durch die Betonung des Zufälligen mit der Wahrscheinlichkeitsrechnung in Beziehung gebracht hatte, sah er ihn doch als selbständig an. In Wirklichkeit aber ist die Kollektivmaßlehre eine Fortbildung der Methoden der Wahrscheinlichkeitsrechnung, deren Ziel ja in der Bestimmung relativer Häufigkeiten besteht, und anwendbar ohne Rücksicht darauf, nach welchem Prinzip der Kollektivgegenstand entstanden ist.

Allgemein sei die Verteilungsfunktion eines beliebigen stetigen Kollektivgegenstandes zum Unterschiede von der speziellen $\varphi(x)$, aber in demselben Sinne verstanden wie diese, bezeichnet mit $\mathfrak{V}(x)$[1]; dann ist

$$\int_a^b \mathfrak{V}(x)\,dx$$

die relative Häufigkeit der Exemplare, deren Argument zwischen a und b liegt, und

$$m\int_a^b \mathfrak{V}(x)\,dx$$

die Anzahl solcher Exemplare, die in einer Kollektivreihe vom Umfange m zu erwarten sind. Insbesondere bedeutet

$$\int_{-\infty}^{X} \mathfrak{V}(x)\,dx$$

mit Rücksicht darauf, daß für außermögliche Argumentwerte $\mathfrak{V}(x) = 0$

1) In der Bezeichnung schließe ich mich der Hauptsache nach an H. Bruns' Werk „Wahrscheinlichkeitsrechnung und Kollektivmaßlehre“, Leipzig 1906, an, das eine bedeutsame Fortbildung dieses Gegenstandes gebracht hat.

ist, die relative Häufigkeit solcher Exemplare, deren Argument *unter* X liegt; ferner ist

$$\int_{-\infty}^{\infty} \mathfrak{V}(x)\,dx = 1. \tag{1}$$

Die Verteilungsfunktion eines unstetigen Gegenstandes werde mit $\mathfrak{U}(x)$ bezeichnet; sind $x_1, x_2, \cdots x_n$ die möglichen Argumentwerte, so ist x nur auf diese angewiesen und daher unstetig; man hat dann in

$$\sum_1^i \mathfrak{U}(x_\nu)$$

die relative Häufigkeit solcher Exemplare, deren Argument höchstens x_i ist, und wiederum

$$\sum_1^n \mathfrak{U}(x_\nu) = 1. \tag{2}$$

Die Beziehung der Verteilungsfunktion zur Verteilungstafel ist in den beiden unterschiedenen Fällen eine ungleiche. Bei einem unstetigen Kollektivgegenstand gibt die primäre Verteilungstafel unmittelbar empirische Werte von $\mathfrak{U}(x)$, indem theoretisch

$$y_i = \mathfrak{U}(x_i) \quad z_i = m\,\mathfrak{U}(x_i) \tag{3}$$

ist. Handelt es sich aber um einen stetigen Kollektivgegenstand, so ist aus der Verteilungstafel ein unmittelbarer Aufschluß über $\mathfrak{V}(x)$ nicht zu erhalten, sondern nur über bestimmte Integrale hiervon, indem theoretisch

$$y_i = \int_{X_i}^{X_{i+1}} \mathfrak{V}(x)\,dx, \qquad z_i = m \int_{X_i}^{X_{i+1}} \mathfrak{V}(x)\,dx \tag{4}$$

ist.

Der Verlauf der Verteilungskurve bei allen bisher untersuchten Kollektivgegenständen kann in großen Zügen dahin gekennzeichnet werden, dass sie im außermöglichen Gebiet mit der Abszissenachse zusammenfällt, von den Enden des möglichen Gebiets gegen einen Gipfelpunkt ansteigt, der zu dem *dichtesten*, relativ häufigsten, Wert des Arguments hinführt. Dabei sind Schwankungen in Folge zufälliger Störungen nicht ausgeschlossen, ja gegen die Enden des möglichen Gebiets die Regel.

179. Summenfunktion. Unter der Summenfunktion soll diejenige Funktion verstanden werden, welche die Summe der Werte der Verteilungsfunktion, beziehungsweise ihr Integral, vom Anfangspunkte des möglichen Argumentgebiets bis zu einem beliebigen möglichen Argumentwert, beziehungsweise bis zu einem beliebigen Wechselpunkt

darstellt, je nachdem es sich um einen unstetigen oder stetigen Kollektivgegenstand handelt. Bezeichnet man sie allgemein mit $\mathfrak{S}(x)$, so ist im ersten Falle

$$\mathfrak{S}(x_i) = \sum_1^i \mathfrak{U}(x_\nu), \tag{5}$$

im zweiten Falle

$$\mathfrak{S}(X_i) = \int_{-\infty}^{X_i} \mathfrak{B}(x)\,dx \tag{6}$$

Die Funktion $\mathfrak{S}(x_i)$ wächst von x_1 an bei Überschreitung jedes möglichen Argumentwerts unstetig, während sie zwischen zwei aufeinander folgenden konstant bleibt, und erreicht bei x_n ihren größten Wert 1, den sie fortan beibehält; ihr geometrisches Bild fällt vor x_1 mit der Abszissenachse zusammen, besteht zwischen x_1 und x_n aus einer Folge staffelförmig angeordneter Strecken parallel zur Achse und jenseits x_n in einer Parallelen im Abstande 1 von der Achse. Aus der Wertreihe der Summenfunktion:

$$\mathfrak{S}(x_1), \quad \mathfrak{S}(x_2), \cdots \quad \mathfrak{S}(x_n)$$

ergibt sich durch Differenzbildung die Wertreihe der Verteilungsfunktion:

$$\mathfrak{U}(x_1) = \mathfrak{S}(x_1), \quad \mathfrak{U}(x_2) = \Delta\,\mathfrak{S}(x_1), \cdots \quad \mathfrak{U}(x_n) = \Delta\,\mathfrak{S}(x_{n-1}).$$

$\mathfrak{S}(X_i)$ wächst vom Anfangspunkt des möglichen Gebiets stetig bis zum Endpunkt desselben, von dem an es den größten Wert 1 beständig beibehält. Kennt man die Wertreihe

$$\mathfrak{S}(X_1) = 0, \quad \mathfrak{S}(X_2), \quad \mathfrak{S}(X_3), \cdots \quad \mathfrak{S}(X_{n+1}),$$

so ergeben sich daraus nicht wie früher Werte der Verteilungsfunktion selbst, sondern bestimmte Integrale derselben, indem

$$\mathfrak{S}(X_2) = \int_{X_1}^{X_2} \mathfrak{B}(x)\,dx, \quad \Delta\,\mathfrak{S}(X_2) = \int_{X_2}^{X_3} \mathfrak{B}(x)\,dx, \quad \cdots \Delta\,\mathfrak{S}(X_n) = \int_{X_n}^{X_{n+1}} \mathfrak{B}(x)\,dx.$$

Beschränkt man sich im Falle der Stetigkeit nicht auf die Wechselpunkte und definiert allgemein

$$\mathfrak{S}(x) = \int_{-\infty}^{x} \mathfrak{B}(x)\,dx, \tag{7}$$

wobei die untere Grenze $-\infty$ zunächst nur formale Bedeutung hat, so folgt daraus

$$\mathfrak{B}(x) = \frac{d\,\mathfrak{S}(x)}{dx}. \tag{8}$$

Rein analytisch aufgefaßt wäre es demnach gleichgültig, welche der beiden Funktionen $\mathfrak{B}(x)$, $\mathfrak{S}(x)$ man direkt bestimmt; denn aus $\mathfrak{B}(x)$ ergibt sich $\mathfrak{S}(x)$ durch Integration, aus $\mathfrak{S}(x)$ erhält man $\mathfrak{B}(x)$ durch Differentiation.

Praktisch hat die Summenfunktion, namentlich bei Voraussetzung der Stetigkeit, den Vorzug eines gleichmäßigeren d. h. von Schwankungen freien Verlaufs, insbesondere aber den Vorzug, daß die Verteilungstafel empirische Werte derselben zu bilden gestattet, was hinsichtlich der Verteilungsfunktion nicht der Fall ist. Die Überlegenheit der Summenfunktion wird sich weiterhin deutlich herausstellen.

180. Durchschnittsbildung. Einem Kollektivgegenstande gegenüber kann eine bestimmte Funktion des Arguments, sie heiße $f(x)$, ein Interesse haben. Der Wertevorrat dieser Funktion hängt von der Verteilung des x, also von der Verteilungsfunktion $\mathfrak{B}(x)$ ab; als Repräsentant dieses Wertvorrats wird der Durchschnitt $\mathfrak{D}[f(x)]$ aller Werte gelten können, und dieser ergibt sich, wenn einmal $\mathfrak{B}(x)$ bekannt ist, nach der Formel

$$\mathfrak{D}[f(x)] = \int_{-\infty}^{\infty} f(x) \mathfrak{B}(x) dx. \tag{9}$$

Wie man erkennt, entspricht diese Gedankenbildung derjenigen des Fehlerrisikos (Nr. 140).

Wählt man $f(x) = x$, so ergibt sich der Durchschnitt der Argumentwerte selbst, ihr arithmetisches Mittel nach Maßgabe ihrer relativen Häufigkeit, kurz der *Argumentdurchschnitt*, in dem man seit jeher eine für die Materie charakteristische Größe erblickt; bezeichnet man ihn mit A, so ist

$$A = \mathfrak{D}(x) = \int_{-\infty}^{\infty} x \mathfrak{B}(x) dx. \tag{10}$$

Mit der Annahme $f(x) = |x - A|$ kommt man zu einer Größe, die im Sinne der Fehlertheorie (Nr. 141) als durchschnittliche Abweichung der Argumentwerte vom arithmetischen Mittel zu bezeichnen wäre; nennt man sie ϑ, so ist

$$\vartheta = \mathfrak{D}[|x - A|] = \int_{-\infty}^{\infty} |x - A| \mathfrak{B}(x) dx. \tag{11}$$

181. Das Exponentialgesetz. Wenn die Verteilungstafel eines Kollektivgegenstandes ein hohes Maß von Symmetrie aufweist, so kann der Versuch gemacht werden, ihr das Exponentialgesetz

$$\varphi(\lambda) = \frac{h}{\sqrt{\pi}} e^{-h^2 \lambda^2}$$

als Verteilungsfunktion zugrunde zu legen; dies der Ausgangspunkt von Fechners Darstellung.

Um diese Supposition an der Tafel prüfen zu können, braucht man den Parameter h; diesen bestimmt Fechner aus der durchschnittlichen Abweichung ϑ nach der Formel (Nr. 141):

$$h = \frac{1}{\vartheta\sqrt{\pi}}$$

und das ϑ selbst empirisch nach der Formel

$$\vartheta = \frac{[|\lambda|]}{m},$$

wobei

$$\lambda_i = -x_i + A, \qquad A = \frac{[xz]}{m} = [xy]$$

und m der Umfang der Kollektivreihe ist.

Ist h berechnet, so hat man in

$$m\int_{X'}^{X''}\varphi(\lambda)d\lambda = m\left\{\int_0^{X''}\varphi(\lambda)d\lambda - \int_0^{X'}\varphi(\lambda)d\lambda\right\} = \frac{m}{2}\left\{\Phi(hX'') - \Phi(hX')\right\}$$

die rechnungsmäßige Anzahl der zwischen die Wechselpunkte X', X'' fallenden Exemplare, die nun mit der beobachteten zu vergleichen ist. Zur Durchführung dieser Rechnungen bedient man sich der Taf. I des Integrals $\Phi(\gamma)$.

Um einer in der Verteilungstafel hervortretenden Asymmetrie beizukommen, wendet Fechner, wie schon bemerkt, das zweiseitige Exponentialgesetz an. Als Ausgangswert dient dann nicht mehr das arithmetische Mittel A, sondern der *dichteste Wert* D, und von diesem aus ist die eben erläuterte Rechnung doppelt, nach beiden Seiten zu führen. Für die Bestimmung von D, das in der Regel in jenes Intervall fällt, zu dem das größte z oder y gehört, entwickelt er besondere Methoden. Alle auf die untere (d. h. mit den kleinen Argumentwerten besetzte) Seite bezüglichen Größen versieht er mit einem untern Akzent ($_,$), alle auf die obere Seite bezüglichen mit einem oberen Akzent ($'$).

Nach diesen Vorbemerkungen wird die folgende Tabelle verständlich sein, die ein Beispiel für Fechners Behandlungsweise von Kollektivgegenständen bietet. Sie betrifft die Rekrutenmaße von 20 Jahrgängen 20-jähriger Leipziger Studenten; der durch die ersten zwei Spalten dargestellten Verteilungstafel ist sowohl das einfache als auch das zweispaltige Exponentialgesetz zugrunde gelegt. Maßeinheit ist der sächs. Zoll = 23,6 *mm*.

Elemente für die Behandlung nach dem einfachen Gauß'schen Gesetz:

$$A = 71{,}75; \quad \vartheta = 2{,}04, \quad m = 2047.$$

Elemente für die Behandlung nach dem zweispaltigen Gesetz:

$$D = 71{,}99; \quad \begin{matrix} \vartheta_{,} = 2{,}16, & m_{,} = 1083{,}5 \\ \vartheta' = 1{,}92, & m' = 963{,}5 \end{matrix}$$

x_i (Zoll)	z_i	Rechnungsmäßige Verteilung (z_i) nach dem einfachen Gesetz	Rechnungsmäßige Verteilung (z_i) nach dem zweispaltigen Gesetz	Differenz zwischen der rechnungsmäß. u. d. beob. Verteil. nach dem einfachen Gesetz	Differenz zwischen der rechnungsmäß. u. d. beob. Verteil. nach dem zweispaltigen Gesetz
60	1	—	—	— 1	— 1
61	0	—	—	0	0
62	0	—	0·5	0	+ 0·5
63	0	1	1·5	+ 1	+ 1·5
64	2	3·5	4	+ 1·5	+ 2
65	15·5	10	12	— 5·5	— 3·5
66	26	26	28	0	+ 2
67	54	58	59	+ 4	+ 5
68	108	110	108	+ 2	0
69	172	179	174	+ 7	+ 2
70	253	252	243	— 1	— 10
71	290	304	298	+ 14	+ 8
72	330·5	315	318	— 15·5	— 12·5
73	296	282	291	— 14	— 5
74	223·5	217	226	— 6·5	+ 2·5
75	142	143	145·5	+ 1	+ 3·5
76	75	81	80·5	+ 6	+ 5·5
77	38	40	37	+ 2	— 1
78	13	17	15	+ 4	+ 2
79	3·5	6	5	+ 2·5	+ 1·5
80	2	2	1	0	— 1
81	1	0·5	—	— 0·5	— 1
82	0·5	—	—	— 0·5	— 0·5
83	0·5	—	—	— 0·5	— 0·5
	2047	2047	2047	abs. 90	abs. 72

Bei der nur schwachen Asymmetrie gibt auch das einfache Exponentialgesetz die Verteilung in nicht unbefriedigender Weise wieder; das zweispaltige paßt sich etwas besser an, die größte Abweichung beträgt dort 15,5, hier nur 12,5, die Summe der absoluten Abweichungen dort 90, hier bloß 72.

182. Reihenentwicklung zur Darstellung willkürlicher Verteilungen. Bei der umfassenden Definition des Kollektivgegenstandes ist es das naturgemäßeste, die Verteilungsfunktion und somit auch die Summenfunktion von vornherein als eine willkürliche Funktion anzusehen, die allen Verteilungen angepaßt werden kann.

Eine solche Funktion ist nur durch einen unendlichen Rechenprozeß darstellbar.[1])

1. Teilt man das Argumentgebiet durch einen Wechselpunkt X in zwei Teile, so ergibt sich nach (1), daß

$$\int_{-\infty}^{X} \mathfrak{W}(x)\,dx + \int_{X}^{\infty} \mathfrak{W}(x)\,dx = 1;$$

mit Benutzung des Diskontinuitätsfaktors sgn $(X-x)$ (lies signum $X-x$), der so definiert ist, daß

$$\operatorname{sgn}(X-x) = 1, \text{ so lange } X-x>0$$
$$\operatorname{sgn}(X-x) = -1, \text{ so lange } X-x<0,$$

verwandelt sich die letzte Gleichung in

$$\int_{-\infty}^{X} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx - \int_{X}^{\infty} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx = 1;$$

da nun

$$\int_{\infty}^{X} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx = \int_{-\infty}^{\infty} \mathfrak{W}(x)\,dx = \mathfrak{S}(X),$$

so ist wegen der vorhergehenden Gleichung

$$\int_{X}^{\infty} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx = \mathfrak{S}(X) - 1,$$

folglich

$$\int_{-\infty}^{X} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx + \int_{X}^{\infty} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx = 2\,\mathfrak{S}(X) - 1$$

und mit Rücksicht auf die Definitionsgleichung (9)

$$2\,\mathfrak{S}(X) - 1 = \int_{-\infty}^{\infty} \operatorname{sgn}(X-x)\,\mathfrak{W}(x)\,dx = \mathfrak{D}[\operatorname{sgn}(X-x)]. \qquad (12)$$

Diese Gleichung wird die Grundlage für die Darstellung von $\mathfrak{S}(X)$ bilden, das ja nach den Ausführungen von Nr. 189 geeignet ist, $\mathfrak{W}(x)$ zu ersetzen.

---

1) Vgl. bezüglich des folgenden *H. Bruns,* Wahrscheinlichkeitsr. u. Kollektivmaßlehre; ferner d. Verf.'s Abhandlung „Die Altersverteilung der Gestorbenen". Mitt. d. Verbandes d. öst.-ung. Privat-Versich.-Anst., Neue Folge, Band 3. Allgemeinere Untersuchungen über die Darstellung willkürlicher Funktionen findet man bei *F. G. Lipps,* Die Theorie der Kollektiv-Gegenstände, Leipzig 1902, S. 89—99.

2. Von den verschiedenen analytischen Darstellungen der diskontinuierlichen Funktion sgn ξ wählen wir die folgende; Es ist

$$\int_{-\infty}^{\infty}\frac{\sin 2\xi u}{u}\,du = \begin{cases} \pi & \text{bei } \xi > 0 \\ 0 & \text{„} \quad \xi = 0 \\ -\pi & \text{„} \quad \xi < 0 \end{cases}$$

infolgedessen

$$\operatorname{sgn}\xi = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin 2\xi u}{u}\,du.$$

Berücksichtigt man, daß

$$\sin 2\xi u = \frac{e^{2\xi u i} - e^{-2\xi u i}}{2i} \qquad (i = \sqrt{-1})$$

und daß

$$\int_{-\infty}^{\infty}\frac{e^{-2\xi u i}}{u}\,du = -\int_{-\infty}^{\infty}\frac{e^{2\xi u i}}{u}\,du,$$

wie man durch Zeichenänderung des u erkennt, so ist weiter

$$\operatorname{sgn}\xi = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{2\xi u i}}{u i}\,du. \tag{13}$$

3. Der Wert des Integrals

$$K = \int_{-\infty}^{\infty} e^{-u^2}\cos 2\xi u\,du$$

ergibt sich durch Differentiation nach dem Parameter ξ, indem

$$\frac{dK}{d\xi} = -\int_{-\infty}^{\infty} 2u e^{-u^2}\sin 2\xi u\,du = \left\{e^{-u^2}\sin 2\xi u\right\}_{-\infty}^{\infty} - 2\xi\int_{-\infty}^{\infty} e^{-u^2}\cos 2\xi u\,du,$$

also

$$\frac{dK}{d\xi} = -2\xi K,$$

folglich

$$K = Ce^{-\xi^2}$$

wird; die Konstante C erhält man durch die Spezialisierung $\xi = 0$, indem dann $K = \int_{-\infty}^{\infty} e^{-u^2}du = \sqrt{\pi}$ wird; somit ist endgültig

$$K = \sqrt{\pi}\,e^{-\xi^2}. \tag{14}$$

Dies leitet auf die in der Wahrscheinlichkeitsrechnung vielfach auftretende Funktion

$$\Phi(\xi) = \frac{2}{\sqrt{\pi}} \int_0^{\xi} e^{-t^2} dt$$

hin; bezeichnet man ihren q-ten Differentialquotienten mit $\Phi(\xi)_q$, so folgt aus einmaliger Differentiation

$$e^{-\xi^2} = \frac{\sqrt{\pi}}{2} \Phi(\xi)_1,$$

und mit Rücksicht auf (14) und die ursprüngliche Bedeutung von K:

$$\Phi(\xi)_1 = \frac{2}{\pi} \int_{-\infty}^{\infty} e^{-u^2} \cos 2\xi u du,$$

woraus durch Integration nach ξ zwischen 0 und ξ

$$\Phi(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-u^2}}{u} \sin 2\xi u du$$

erhalten wird.

Ersetzt man wie oben $\sin 2\xi u$ durch

$$\frac{e^{2\xi u i} - e^{-2\xi u i}}{2i}$$

und beachtet die durch Zeichenänderung des u erweisbare Relation

$$\int_{-\infty}^{\infty} \frac{e^{-u^2 - 2\xi u i}}{u} du = -\int_{-\infty}^{\infty} \frac{e^{-u^2 + 2\xi u i}}{u} du,$$

so kann auch

$$\Phi(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-u^2 + 2\xi u i}}{u i} du$$

geschrieben werden, woraus sich unmittelbar eine Darstellung von $\Phi(\xi)_q$ ableitet, indem

$$\Phi(\xi)_q = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-u^2 + 2\xi u i}}{u i} (2ui)^q du \tag{15}$$

wird.

4. Die Exponentialfunktion $e^{-u^2 - 2\xi u}$ gestattet eine durchaus konvergente Entwicklung nach Potenzen von u; ihre Koeffizienten werden rationale Funktionen von ξ sein, sie möge daher in der Form

$$e^{-u^2 - 2\xi u} = \mathfrak{R}(\xi)_0 + \mathfrak{R}(\xi)_1 2u + \mathfrak{R}(\xi)_2 (2u)^2 + \cdots = \sum_0^{\infty} \mathfrak{R}(\xi)_q (2u)^q \tag{16}$$

angesetzt werden. Andererseits ist, wenn man jeden Faktor für sich entwickelt:

$$e^{-u^2-2\xi u} = \left(\frac{1}{0!}-\frac{u^2}{1!}+\frac{u^4}{2!}-\frac{u^6}{3!}+\cdots\right)\left(\frac{1}{0!}-\frac{2\xi u}{1!}+\frac{(2\xi u)^2}{2!}-\frac{(2\xi u)^3}{3!}+\cdots\right);$$

durch Vergleichung der Koeffizienten gleicher Potenzen von u ergibt sich daraus folgende Darstellung der $\mathfrak{R}$:

$$\begin{aligned}
\mathfrak{R}(\xi)_0 &= 1\\
2\mathfrak{R}(\xi)_1 &= \frac{-2\xi}{0!\,1!}\\
2^2\mathfrak{R}(\xi)_2 &= \frac{(2\xi)^2}{0!\,2!}-\frac{1}{1!\,0!}\\
2^3\mathfrak{R}(\xi)_3 &= -\frac{(2\xi)^3}{0!\,3!}+\frac{2\xi}{1!\,1!}\\
2^4\mathfrak{R}(\xi)_4 &= \frac{(2\xi)^4}{0!\,4!}-\frac{(2\xi)^2}{1!\,2!}+\frac{1}{2!\,0!},
\end{aligned} \tag{17}$$

allgemein:

$$\begin{aligned}
2^{2p}\mathfrak{R}(\xi)_{2p} &= \frac{(2\xi)^{2p}}{0!\,\overline{2p}!}-\frac{(2\xi)^{2p-2}}{1!\,\overline{2p-2}!}+\frac{(2\xi)^{2p-4}}{2!\,\overline{2p-4}!}+\cdots+\frac{(-1)^p}{p!\,0!}\\
2^{2p+1}\mathfrak{R}(\xi)_{2p+1} &= -\frac{(2\xi)^{2p+1}}{0!\,\overline{2p+1}!}+\frac{(2\xi)^{2p-1}}{1!\,\overline{2p-1}!}-\frac{(2\xi)^{2p-3}}{2!\,\overline{2p-3}!}+\cdots+\frac{(-1)^{p+1}2\xi}{p!\,1!}.
\end{aligned} \tag{18}$$

5. Wir nehmen nun die in (13) gefundene Darstellung von $\operatorname{sgn}\xi$ wieder auf und schreiben sie mit dem Argument $X-x$; dies führt zu

$$\begin{aligned}
\operatorname{sgn}(X-x) &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{2Xui-2xui}}{ui}\,du\\
&= \frac{1}{\pi}\int_{0}^{\infty}\frac{e^{2Xui-u^2}}{ui}\cdot e^{u^2-2xui}\,du;
\end{aligned}$$

mit Benutzung von (16) ergibt sich aber

$$e^{-u^2-2xui} = \sum_0^\infty{}_q\,\mathfrak{R}(x)_q(2ui)^q,$$

folglich ist weiter

$$\operatorname{sgn}(X-x) = \frac{1}{\pi}\sum_0^\infty{}_q\,\mathfrak{R}(x)_q\int_{-\infty}^{\infty}\frac{e^{-u^2+2Xui}}{ui}(2ui)^q\,du$$

demnach mit Rücksicht auf (15):

$$\operatorname{sgn}(X-x) = \sum_0^\infty{}_q\,\mathfrak{R}(x)_q\,\Phi(X)_q.$$

Es erweist sich als zweckmäßig, statt mit den Argumentwerten selbst mit ihren Abweichungen von einem Ausgangswert a zu rechnen; führt man dementsprechend, das a vorläufig unbestimmt lassend, unter Verwendung einer ebenfalls noch frei wählbaren *positiven* Konstante h die neue Variable

$$v = h(x - a) \tag{19}$$

ein, so erreicht man damit auch den Vorteil, daß man durch passende Wahl von a und h auf eine Vereinfachung hinwirken kann; es soll dann entsprechend

$$V = h(X - a) \tag{19*}$$

sein. Da nun $V - v = h(X - x)$ dasselbe Vorzeichen hat wie $X - x$, so besteht auch der Ansatz:

$$\operatorname{sgn}(V - v) = \sum_0^\infty {}_q \mathfrak{R}(v)_q \Phi(V)_q.$$

Die Einführung dieses Ausdruckes in (12), wobei zu beachten ist, daß sich die $\mathfrak{D}$-Operation lediglich auf das von x abhängige $\mathfrak{R}(v)_q$ erstreckt, liefert das *Endresultat*:

$$2\mathfrak{S}(X) - 1 = \sum_0^\infty {}_q \mathfrak{D}[\mathfrak{R}(v)_q] \Phi(V)_q. \tag{20}$$

Dies ist die von Bruns entwickelte Φ-Reihe zur Darstellung beliebiger Verteilungen. Ihre praktische Verwendbarkeit hängt ab von der Raschheit der Abnahme der Glieder. Noch bleibt die Herstellung der symbolisch geschriebenen Koeffizienten $\mathfrak{D}[\mathfrak{R}(v)_q]$ übrig. Doch wird diese besser zu verbinden sein mit dem Rechnungsschematismus für die Durchführung des ganzen Verfahrens.

§ 2. Methodische Durchführung der Rechnungen.

183. Die Koeffizienten der Φ-Reihe. Da die $\mathfrak{R}(\xi)_q$ laut (17) nach ξ geordnete Polynome sind, so werden die $\mathfrak{R}(v)_q = \mathfrak{R}\{h(x-a)\}_q$ nach Potenzen von $x - a$ fortschreiten und es wird daher bei der Durchschnittsbildung auf die Durchschnittswerte der verschiedenen Potenzen der Abweichungen vom Ausgangswerte ankommen; wir setzen allgemein

$$\eta_p^p = \mathfrak{D}[(x - a^p)] = \int_{-\infty}^{\infty} (x - a)^p \mathfrak{B}(x) dx; \tag{21}$$

dann ist insbesondere

$$\eta_1 = \mathfrak{D}[x - a] = \mathfrak{D}(x) - a = A - a; \tag{22}$$

es wird also $\eta_1 = 0$, wenn man den Argumentdurchschnitt A als

Ausgangswert wählt; in bezug auf diesen werde konsequent μ statt η geschrieben, so daß

$$\mu_p^p = \mathfrak{D}[(x-A)^p] = \int_{-\infty}^{\infty}(x-A)^p \mathfrak{W}(x)\,dx. \tag{23}$$

Dies vorausgeschickt und bei Festhaltung an dieser Wahl des Ausgangswertes wird auf Grund von (17):

$$\mathfrak{D}[\mathfrak{R}(v)_0] = 1,$$
$$2\mathfrak{D}[\mathfrak{R}(v)_1] = -2h\mathfrak{D}(x-A) = 0,$$
$$2^2\mathfrak{D}[\mathfrak{R}(v)_2] = \frac{4h^2}{2}\mathfrak{D}[(x-A)^2] - 1 = 2h^2\mu_2^2 - 1;$$

verfügt man also über die noch unbestimmte Zahl h derart, daß

$$2h^2\mu_2^2 - 1 = 0,$$

also

$$h = \frac{1}{\mu_2\sqrt{2}} \tag{24}$$

wird, so fällt auch das mit $\mathfrak{R}(v)_2$ behaftete Glied aus.

Die Größe μ_2 als Wurzel aus dem Durchschnitt der Quadrate der Abweichungen vom arithmetischen Mittel entspricht dem mittleren Fehler in der Fehlertheorie und der Parameter h, wie die Formel (24) zeigt (vgl. mit Nr. 142), dem Präzisionsmaß. Bruns hat μ_2 unter der Bezeichnung „Streuung von x, str (x)“ als charakteristisches Element eingeführt, das ein Urteil über die *Ausbreitung* der Kollektivreihe gestattet. Sowie nur eine sehr geringe Wahrscheinlichkeit dafür besteht, daß der Fehler einer Beobachtung den dreifachen mittleren Fehler der betreffenden Beobachtungsart überschreiten werde, so ist auch anzunehmen, daß die Ausbreitung des Argumentgebiets eines Kollektivgegenstandes die sechsfache Streuung nicht übertreffe, anders gesagt, daß die Hauptmasse der Exemplare in diesem Intervall liege. Die Formel (24) kann hiernach auch so geschrieben werden:

$$h = \frac{1}{\operatorname{str}(x)\sqrt{2}}. \tag{24*}$$

In weiterer Verfolgung der Gleichungen (17) und unter steter Berücksichtigung der obigen Festsetzungen hat man nun:

$$2^3\mathfrak{D}[\mathfrak{R}(v)_3] = -\frac{8h^3\mu_3^3}{6} + 2h\mu_1 = -\frac{4h^3\mu_3^3}{3}$$

$$2^4\mathfrak{D}[\mathfrak{R}(v)_4] = \frac{16h^4\mu_4^4}{24} - \frac{4h^2\mu_2^2}{2} + \frac{1}{2} = \frac{2h^4\mu_4^4}{3} - \frac{1}{2}$$

$$2^5\mathfrak{D}[\mathfrak{R}(v)_5] = -\frac{32h^5\mu_5^5}{120} + \frac{8h^3\mu_3^3}{6} - \frac{2h\mu_1}{2} = -\frac{4h^5\mu_5^5}{15} + \frac{4h^3\mu_3^3}{3}$$

$$2^6\mathfrak{D}[\mathfrak{R}(v)_6] = \frac{64h^6\mu_6^6}{720} - \frac{16h^4\mu_4^4}{24} + \frac{4h^2\mu_2^2}{4} - \frac{1}{6} = \frac{4h^6\mu_6^6}{45} - \frac{2h^4\mu_4^4}{3} + \frac{1}{3};$$

wird also zur Abkürzung gesetzt:

$$
\begin{aligned}
D_3 &= -\frac{2h^3\mu_3^3}{3} \\
D_4 &= \frac{h^4\mu_4^4}{3} - \frac{1}{4} \\
D_5 &= -\frac{2h_5\mu_5^5}{15} + \frac{2h^3\mu_3^3}{3} \\
D_6 &= \frac{2h^6\mu_6^6}{45} - \frac{h^4\mu_4^4}{3} + \frac{1}{6},
\end{aligned} \tag{25}
$$

so schreibt sich die Formel (20) endgültig:

$$
\begin{aligned}
2\mathfrak{S}(X) - 1 = \Phi(V) + D_3 \frac{\Phi(V)_3}{2^2} + D_4 \frac{\Phi(V)_4}{2^3} \\
+ D_5 \frac{\Phi(V)_5}{2^4} + D_6 \frac{\Phi(V)_6}{2^5} + \cdots
\end{aligned} \tag{26}
$$

Bruns nennt sie die *Normalform* der Φ-Reihe mit Rücksicht auf die Vereinfachungen, die sie durch die Wahl von a und h erfahren hat.

Die Auswertung dieser Formel erfordert als beschwerlichsten Teil der Arbeit die Ausrechnung der Elemente $D_3, D_4 \ldots$ auf Grund der Verteilungstafel, außerdem Tabellen der Werte von

$$\Phi(V); \quad \frac{\Phi(V)_3}{2^2}, \quad \frac{\Phi(V)_4}{2^3}, \ldots;$$

für $\Phi(V)$ sind solche längst vorhanden; dem Buche ist unter I. eine siebenstellige Tafel dieser Funktion, unter II. eine vierstellige, nach einem kleineren Argumentintervalle fortschreitende Tafel beigegeben. Für die folgenden Werte bis einschließlich zur Ordnung 6 hat Bruns Tabellen berechnet, die unter III. dem Buche angeschlossen sind.[1])

Beschränkt man die Formel (26) auf das erste Glied, so gibt sie

$$\mathfrak{S}(X) = \frac{1}{2} + \frac{1}{2}\Phi(V) = \frac{1}{2} + \frac{1}{\sqrt{\pi}}\int_0^V e^{-t^2}dt = \frac{1}{\sqrt{\pi}}\int_{-\infty}^V e^{-t^2}dt$$

entsprechend dem Gaußschen Exponentialgesetz. Die Fortsetzung der Formel bringt also die Abweichungen von diesem speziellen Gesetz zum Ausdruck; je geringer diese Abweichungen, mit umso weniger Gliedern wird man das Auslangen finden.

Hat man $\mathfrak{S}(X)$ mit Beschränkung auf eine bestimmte Anzahl von Gliedern numerisch dargestellt, dann bleibt zu prüfen, in welchem Ausmaße sich diese Darstellung der Verteilungstafel anpaßt. Zu diesem Zwecke berechnet man die den Wechselpunkten

$$X_1, X_2, \cdots X_{n+1}$$

---

1) Mit freundlicher Erlaubnis des Autors.

zugeordneten Werte des Hilfsarguments $v = h(x - A)$:

$$V_1, V_2, \cdots V_{n+1},$$

bestimmt mit Hilfe derselben

$$\mathfrak{S}(X_1), \mathfrak{S}(X_2), \cdots \mathfrak{S}(X_{n+1}),$$

bildet die Differenzenreihe

$$\mathfrak{y}_1, \mathfrak{y}_2, \cdots \mathfrak{y}_n$$

und vergleicht diese mit der Zahlenfolge

$$y_1, y_2, \cdots y_n$$

oder aber die Wertreihe

$$m\mathfrak{y}_1, m\mathfrak{y}_2, \cdots m\mathfrak{y}_n$$

mit

$$z_1, z_2, \cdots z_n.$$

184. Summenverfahren bei freier Wahl des Ausgangswertes. Die Bestimmung der D erfordert die Kenntnis von A und der μ_p^p von $p = 2$ angefangen; denn μ_2^2 bestimmt die Streuung und h.

Was A betrifft, so ergibt sich die praktische Formel dafür aus der theoretischen (10):

$$A = \int_{-\infty}^{\infty} x\mathfrak{B}(x)dx,$$

indem man das Integrationsgebiet durch die Wechselpunkte in Teilintervalle zerlegt und in jedem solchen das variable x durch den mittleren Wert ersetzt; so wird

$$A = \sum_1^n x_i \int_{X_i}^{X_{i+1}} \mathfrak{B}(x)dx = \sum x_i y_i = \frac{\sum x_i z_i}{m}. \tag{27}$$

Nach dem gleichen Vorgange erhält man aus (23):

$$\mu_p^p = \sum_1^n (x_i - A)^p \int_{X_i}^{X_{i+1}} \mathfrak{B}(x)dx = \sum (x_i - A)^p y_i = \frac{\sum (x_i - A)^p z_i}{m}. \tag{28}$$

Das unmittelbare Rechnen nach diesen Formeln gestaltet sich schon bei einer mäßig ausgedehnten Verteilungstafel wegen der vielen Potenzierungen und Multiplikationen sehr beschwerlich; bei Formel (28) kommt noch der Umstand erschwerend hinzu, daß A fast ausnahmslos in das Innere eines Intervalls fallen und deshalb zu unbequemen Zahlen $x - A$ führen wird.

Der letztern Schwierigkeit wird dadurch abgeholfen, daß man statt A einen der in der Tafel angeführten Argumentwerte selbst, er heiße a, zum Ausgangspunkt nimmt; dadurch werden sämtliche Differenzen $x_i - a$ Vielfache der Intervallgröße δ, und es treten an ihre Stelle nach Absonderung des δ *ganze* Zahlen. Zum Schlusse bleibt der Übergang von a zu A zu vollziehen.

Sind

$$x_1,\ x_2,\ \cdots x_k,\ \cdots x_{n-1},\ x_n$$

die äquidistanten Argumentwerte der Tafel und trifft man die Wahl

$$x_k = a,$$

so wird

$$x_{k-\nu} - a = -\nu\delta,\quad x_{k+\nu} - a = \nu\delta.$$

Zur Berechnung von

$$\eta_p^p = \frac{\sum_1^n (x_i - a)^p z_i}{m} \tag{29}$$

hat man für ein *ungerades* p:

$$\sum_1^n (x_i - a)^p z_i = -\delta^p \sum_1^{k-1} \nu^p z_{k-\nu} + \delta^p \sum_{n-k}^{1} \nu^p z_{k+\nu}, \tag{30}$$

für ein *gerades* p:

$$\sum_1^n (x_i - a)^p z_i = \quad \delta^p \sum_1^{k-1} \nu^p z_{k-\nu} + \delta^p \sum_{n-k}^{1} \nu^p z_{k+\nu}, \tag{31}$$

so daß es nur noch auf die Bildung der Summen

$$\sum_1^{k-1} \nu^p z_{k-\nu} = 1^p z_{k-1} + 2^p z_{k-2} + \cdots + (k-1)^p z_1 \tag{32}$$

$$\sum_{n-k}^{1} \nu^p z_{k+\nu} = (n-k)^p z_n + (n-k-1)^p z_{n-1} + \cdots + 1^p z_{k+1} \tag{33}$$

ankommt; diese aber lassen sich bei jedem p durch ein mechanisches Verfahren, das *nur* Additionen erfordert und daher als *Summen*verfahren bezeichnet wird, berechnen.

Dieses Verfahren, auf den „oberen" Teil der Tafel, also auf jenen mit den Argumentwerten $x_1,\ x_2,\ \cdots x_{k-1}$ angewendet, besteht in folgendem. Man bildet durch sukzessives Addieren nach und nach die *ersten Partialsummen:*

$$\begin{aligned}
s_1^{(1)} & & &= z_1 \\
s_2^{(1)} &= s_1^{(1)} &+ z_2 &= z_1 + z_2 \\
s_3^{(1)} &= s_2^{(1)} &+ z_3 &= z_1 + z_2 + z_3 \\
&\ldots \\
s_{k-2}^{(1)} &= s_{k-3}^{(1)} &+ z_{k-2} &= z_1 + z_2 + z_3 + \cdots + z_{k-2},
\end{aligned}$$

und aus diesen die *erste Totalsumme:*

$$\bar{S}_1 = (k-2)\,z_1 + (k-3)\,z_2 + \cdots + z_{k-2};$$

fügt man dazu die *nullte Totalsumme*

$$\bar{S}_0 = s_{k-2}^{(1)} + z_{k-1} = z_1 + z_2 + \cdots + z_{k-1} \tag{34}$$

hinzu, so entsteht

$$\bar{S}_1 S + {}_0^{-} = \sum_1^{k-1} \nu z_{k-\nu}, \tag{35}$$

also die erste der Summen (32).

Hierauf bilde man aus den ersten Partialsummen nach demselben Prinzip, nach welchem sie selbst aus den z gebildet wurden, die *zweiten:*

$$\begin{aligned}
s_1^{(2)} &= s_1^{(1)} & &= z_1 \\
s_2^{(2)} &= s_1^{(2)} &+ s_2^{(1)} &= 2z_1 + z_2 \\
s_3^{(2)} &= s_2^{(2)} &+ s_3^{(1)} &= 3z_1 + 2z_2 + z_3 \\
&\ldots \\
s_{k-3}^{(2)} &= s_{k-4}^{(2)} &+ s_{k-3}^{(1)} &= (k-3)z_1 + (k-4)z_2 + (k-5)z_3 + \cdots + z_{k-3},
\end{aligned}$$

und aus ihnen die *zweite Totalsumme:*

$$\bar{S}_2 = \frac{(k-2)(k-3)}{2} z_1 + \frac{(k-3)(k-4)}{2} z_2 + \cdots + z_{k-3},$$

woraus

$$2\,\bar{S}_2 = (k-2)(k-3)z_1 + (k-3)(k-4)z_2 + \cdots + 2z_{k-3};$$

das allgemeine Glied der rechten Seite, $(k-j-1)(k-j-2)z_j$, kann entwickelt werden in

$$(k-j)^2 z_j - 3\,(k-j) z_j + 2 z_j;$$

infolgedessen hat man

$$2\,\bar{S}_2 = \sum_3^{k-1} \nu^2 z_{k-\nu} - 3 \sum_3^{k-1} \nu z_{k-\nu} + 2 \sum_3^{k-1} z_{k-\nu};$$

um nun rechts die Summen von 1 an laufen zu lassen, hat man

$$2^2 z_{k-2} + 1^2 z_{k-1} - 3\,(2 z_{k-2} + 1 z_{k-1}) + 2\,(z_{k-2} + z_{k-1})$$

hinzuzufügen; da sich dies aber auf Null reduziert, so ist auch

$$2 S_2^- = \sum_1^{k-1} \nu^2 z_{k-\nu} - 3 \sum_1^{k-1} \nu z_{k-\nu} + 2 \sum_1^{k-1} z_{k-\nu},$$

mit Rücksicht auf (34) und (35) also

$$2 S_2^- = \sum_1^{k-1} \nu^2 z_{k-\nu} - 3 (S_1^- + S_0^-) + 2 S_0^-,$$

woraus

$$\sum_1^{k-1} \nu^2 z_{k-\nu} = 2 S_2^- + 3 S_1^- + S_0^-, \tag{36}$$

und damit ist die zweite der Summen (32) gefunden.

Die *dritten Partialsummen*

$$\begin{aligned}
s_1^{(3)} &= s_1^{(2)} && = z_1 \\
s_2^{(3)} &= s_1^{(3)} + s_2^{(2)} && = 3 z_1 + z_2 \\
s_3^{(3)} &= s_2^{(3)} + s_3^{(2)} && = 6 z_1 + 3 z_2 + z_1 \\
& \ldots \\
s_{k-4}^{(3)} &= s_{k-5}^{(3)} + s_{k-4}^{(2)} && = \frac{(k-3)(k-4)}{2} z_1 + \frac{(k-4)(k-5)}{2} z_2 + \cdots + z_{k-4}
\end{aligned}$$

geben die *dritte Totalsumme*

$$S_3^- = \frac{(k-2)(k-3)(k-4)}{1 \cdot 2 \cdot 3} z_1 + \frac{(k-3)(k-4)(k-5)}{1 \cdot 2 \cdot 3} z_2 + \cdots + z_{k-4},$$

woraus

$$6 S_3^- = (k-2)(k-3)(k-4) z_1 + (k-3)(k-4)(k-5) z_2 + \cdots + 6 z_{k-4};$$

das allgemeine Glied $(k-j-1)(k-j-2)(k-j-3) z_j$ läßt sich entwickeln in

$$(k-j)^3 z_j - 6 (k-j)^2 z_j + 11 (k-j) z_j - 6 z_j,$$

demnach ist

$$6 S_3^- = \sum_4^{k-1} \nu^3 z_{k-\nu} - 6 \sum_4^{k-1} \nu^2 z_{k-\nu} + 11 \sum_4^{k-1} \nu z_{k-\nu} - 6 \sum_4^{k-1} z_{k-\nu};$$

sollen die Summen von 1 an laufen, so hat man hinzuzufügen

$$\sum_1^{3} \nu^3 z_{k-\nu} - 6 \sum_1^{3} \nu^2 z_{k-\nu} + 11 \sum_1^{3} \nu z_{k-\nu} - 6 \sum_1^{3} z_{k-\nu};$$

da sich dies aber auf Null reduziert, so ist auch

$$6 S_3^- = \sum_1^{k-1} \nu^3 z_{k-\nu} - 6 \sum_1^{k-1} \nu^2 z_{k-\nu} + 11 \sum_1^{k-1} \nu z_{k-\nu} - 6 \sum_1^{k-1} z_{k-\nu},$$

mit Beziehung auf (34), (35), (36) weiter

$$6S_3^- = \sum_1^{k-1} \nu^3 z_{k-\nu} - 6(2S_2^- + 3S_1^- + S_0^-) + 11(S_1^- + S_0^-) - 6S_0^-,$$

woraus sich die dritte der Summen (32) ergibt:

$$\sum_1^{k-1} \nu^3 z_{k-\nu} = 6S_3^- + 12S_2^- + 7S_1^- + S_0^-. \tag{37}$$

Die vierte Summe erhält man auf analogem Wege, sie hat den Ausdruck

$$\sum_1^{k-1} \nu^4 k_{s-\nu} = 24S_4^- + 60S_3^- + 50S_2^- + 15S_1^- + S_0^-. \tag{38}$$

Bezüglich der Summen (33) genügt die Bemerkung, daß sie durch eine vom untern Ende der Tafel gegen x_k hin geführte ebenso angelegte Rechnung gefunden werden; die Totalsummen, auf die man dabei kommt, seien mit S_0^+, S_1^+, S_2^+, $\cdots$ bezeichnet.

Trägt man dann all dies in (30) und (31) ein und bedient man sich dabei zur Vereinfachung der Formeln der Abkürzungen

$$\begin{aligned} S_i^+ + S_i^- &= \Sigma_i \\ S_i^+ - S_i^- &= \Delta_i \end{aligned} \tag{39}$$

so ergeben sich im Sinne von (34) und (29) die folgenden Resultate:

$$\begin{aligned} m &= \Sigma_0 + z_k \\ m\eta_1 &= \delta\ (\Delta_1 + \Delta_0) \\ m\eta_2^2 &= \delta^2 (2\,\Sigma_2 + 3\,\Sigma_1 + \Sigma_0) \\ m\eta_3^3 &= \delta^3 (6\,\Delta_3 + 12\,\Delta_2 + 7\,\Delta_1 + \Delta_0) \\ m\eta_4^4 &= \delta^4 (24\,\Sigma_4 + 60\,\Sigma_3 + 50\,\Sigma_2 + 15\,\Sigma_1 + \Sigma_0)\,[1]. \end{aligned} \tag{40}$$

Über die Wahl des Ausgangswerts $x_k = a$ ist bisher nichts bemerkt, sie ist im wesentlichen freigestellt. Seine Verlegung in etwa die Mitte der Tafel, wodurch diese in zwei nahezu gleiche Teile zerfällt, einen obern und einen untern, hat den Vorteil, daß die Summenbildung in zwei nahezu gleichmäßigen Partien verläuft und nicht zu allzu großen Zahlen führt, was bei einheitlicher Rechnungsanlage und ausgedehnter Verteilungstafel immer eintreten würde; daher ist nur bei kleinen Tafeln die Verlegung von a an das Ende der Tafel ratsam.

1) Ohne auf die allgemeine Darstellung der Koeffizienten einzugehen, mögen diese auch noch für die zwei nächsten Ordnungen angegeben werden:

für $p = 5$: 120, 360, 390, 180, 31, 1;

„ $p = 6$: 720, 2520, 3360, 2100, 602, 63, 1.

Das Schema der Summenbildung bis zur vierten Ordnung gestaltet sich wie folgt:

Verteilungstafel			Summentafel			
i	x_i	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$	$s_i^{(4)}$
1	x_1	z_1	$s_1^{(1)}$	$s_1^{(2)}$	$s_1^{(3)}$	$s_1^{(4)}$
2	x_2	z_2	$s_2^{(1)}$	$s_2^{(2)}$	$s_2^{(3)}$	.
3	x_3	z_3	$s_3^{(1)}$	$s_3^{(2)}$	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	$s_{k-5}^{(4)}$
.	.	.	.	.	$s_{k-4}^{(3)}$	S_4^-
.	.	.	.	$s_{k-3}^{(2)}$	S_3^-	
$k-1$	x_{k-2}	z_{k-2}	$s_{k-2}^{(1)}$	S_2^-		
$k-2$	x_{k-1}	z_{k-1}	S_1^-			
k	$x_k = a$	z_k	S_1^+			
$k+1$	x_{k+1}	z_{k+1}	$s_{k+2}^{(1)}$	S_2^+		
$k+2$	x_{k+2}	z_{k+2}	.	$s_{k+3}^{(2)}$	S_3^+	
.	.	.	.	.	$s_{k+4}^{(3)}$	S_4^+
.	.	.	.	.	.	$s_{k+5}^{(4)}$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
$n-2$	x_{n-2}	z_{n-2}	$s_{n-2}^{(1)}$	$s_{n-2}^{(2)}$	.	.
$n-1$	x_{n-1}	z_{n-1}	$s_{n-1}^{(1)}$	$s_{n-1}^{(2)}$	$s_{n-1}^{(3)}$	.
n	x_n	z_n	$s_n^{(1)}$	$s_n^{(2)}$	$s_n^{(3)}$	$s_n^{(4)}$

Eine *Kontrolle* ergibt sich daraus, daß nach der Bildungsweise der Partial- und Totalsummen die folgenden Relationen bestehen:

$$z_{k-1} + s_{k-2}^{(1)} = S_0^-$$
$$s_{k-2}^{(1)} + s_{k-3}^{(2)} = S_1^-$$
$$s_{k-3}^{(2)} + s_{k-4}^{(3)} = S_2^-$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

und analog im untern Teil der Tafel.

185. Übergang zum Argumentdurchschnitt als Ausgangswert. Zur Durchführung der Formel (26) braucht man allgemein:

$$\mu_p^p = \mathfrak{D}\,[(x-A)^p],$$

während das Summenverfahren zu

$$\eta_p^p = \mathfrak{D}\,[(x-a)^p]$$

geführt hat. Nun ist vermöge der Gleichung (22)

$$x - A = x - a - (A - a) = x - a - \eta_1,$$

mithin

$$(x - A)^p = (x - a)^p - \binom{p}{1}(x - a)^{p-1}\eta_1 + \binom{p}{2}(x - a)^{p-2}\eta_1^2 + \cdots + (-1)^{p-1}\binom{p}{1}(x - a)\eta_1^{p-1} + (-1)^p\eta_1^p;$$

daraus ergibt sich:

$$\mu_p^p = \eta_p^p - \binom{p}{1}\eta_{p-1}^{p-1}\eta_1 + \binom{p}{2}\eta_{p-2}^{p-2}\eta_1^2 - \cdots + (-1)^{p-1}(p-1)\eta_1^p, \quad (42)$$

wobei das letzte Glied durch Zusammenfassung der beiden früheren Schlußglieder entstanden ist. Man hat also für $p = 1, 2, 3, 4$:

$$\begin{aligned} \mu_1 &= 0 \\ \mu_2^2 &= \eta_2^2 - \eta_1^2 \\ \mu_3^3 &= \eta_3^3 - 3\eta_2^2\eta_1 + 2\eta_1^3 \\ \mu_4^4 &= \eta_4^4 - 4\eta_3^3\eta_1 + 6\eta_2^2\eta_1^2 - 3\eta_1^4. \end{aligned} \quad (43)$$

Hiermit sind alle Elemente bestimmt, die zur Durchführung der Formel (26) erfordert werden[1]).

§ 3. Anwendungen.

186. Beispiel LXVII. Verteilung der Endnullen in einer Logarithmentafel. Als Beispiel eines in mehrfacher Beziehung bemerkenswerten unstetigen Kollektivgegenstandes benutzen wir das Resultat der Auszählung der Endnullen in 1000 Spalten von Vegas Thesaurus Logarithmorum, welches Bruns bekannt gemacht hat.[2]) Da die Logarithmen den Zahlen durch ein arithmetisches Gesetz zugeordnet sind, so ist auch jede einzelne Stelle der Mantisse in gesetzmäßiger Weise bestimmt; es kann daher nicht als zufällig bezeichnet werden, ob an der Endstelle des Logarithmus einer gewissen Zahl Null oder irgend eine andere Ziffer steht; im Sinne der Fechnerschen Definition würde es sich also hier nicht um eine Beobachtungsreihe handeln, die den Kollektivgegenständen zuzuzählen ist; nach der allgemeinen Auffassung, die hier vertreten ist, handelt es sich um einen solchen, und zwar sind seine Exemplare die einzelnen Spalten (mit je 60 Mantissen), der Umfang also 1000, das Argument x die in der Spalte angetroffene Anzahl von Endnullen. Das mögliche Gebiet des Arguments umfaßt die ganzen Zahlen von 0 bis 60, da auch solche Spalten *denkbar* sind, in denen keine Endnull vorkommt, und solche,

---

1) Vgl. zu den beiden letzten Nrn. F. G. Lipps, Theorie der Kollektivgegenst., p. 178—189.

2) a. a. O., p. 256.

in denen jede Mantisse mit Null endet. Die Auszählung ergab keine Spalte ohne Endnull und keine mit mehr als 14 Endnullen; würde sie auf mehr, etwa auf die doppelte Zahl von Spalten, erstreckt werden, so könnten ganz wohl diese Grenzen weiter hinausrücken.

Die folgende Tabelle enthält in den ersten zwei Kolonnen die (primäre) Verteilungstafel, in den nächsten zwei die daraus abgeleiteten empirischen Werte der Verteilungsfunktion $\mathfrak{U}(x)$ und der Summenfunktion $\mathfrak{S}(x)$, in den letzten zwei die Summenrechnung, soweit sie zur Ermittlung des Argumentdurchschnitts und der Streuung erforderlich ist.

x_i	z_i	$\mathfrak{U}(x_i) = \frac{z_i}{m}$	$\sum_1^i \mathfrak{U}(x_i) = \mathfrak{S}(x_i)$	$s_i^{(1)}$	$s_i^{(2)}$
1	6	0,006	0,006	6	6
2	36	0,036	0,042	42	48
3	78	0,078	0,120	120	168
4	149	0,149	0,269	269	437
5	161	0,161	0,430	430	867
6	183	0,183	0,613	613	1480
7	134	0,134	0,747	747	2227
8	114	0,114	0,861	861	3088
9	74	0,074	0,935	935	4023
10	34	0,034	0,969	969	4992
11	19	0,019	0,988	988	5980
12	10	0,010	0,998	998	23316
13	0	0,000	0,998	6978	
14	2	0,002	1,000		
	1000				

Der Tafel ist beispielsweise zu entnehmen, daß die relative Häufigkeit von Spalten mit 7 Endnullen dieser Erfahrung zufolge 0,134, die relative Häufigkeit von Spalten mit *höchstens* 7 Endnullen 0,747 beträgt, daß also die Wahrscheinlichkeit, eine beliebige Spalte werde 7 bzw. höchstens 7 Endnullen enthalten, nahe mit 0,134 bzw. 0,747 oder $\frac{3}{4}$ zu bewerten sei.

Die Summenrechnung ist in einem Zuge durchgeführt, folglich $a = 14$ als Ausgangswert genommen; es existieren also nur Totalsummen S^-, so daß laut (39) $\Sigma_i = S_i^-$, $\Delta_i = - S_i^-$ zu nehmen ist. Die Tabelle liefert nun:

$$S_0^- = 998 + 0 = 998, \qquad m = S_0^- + z_{14} = 1000;$$
$$S_1^- = 6978,$$
$$S_2^- = 23316,$$

hieraus ergibt sich mittels (40) und (22):

$$1000\eta_1 = -6978 - 998 = -7976, \quad \eta_1 = -7{,}976,$$

$$A = 14 - 7{,}976 = 6{,}024,$$

$$1000\,\eta_2^2 = 46632 + 20934 + 998 = 68564, \quad \eta_2^2 = 68{,}564,$$

weiter nach (43):

$$\mu_2^2 = 68{,}564 - 63{,}617 = 4{,}947,$$

$$\mathrm{str}\,(x) = 2{,}224.$$

Diesen Resultaten ist weiter zu entnehmen, daß die durchschnittliche Zahl der in einer Spalte auftretenden Endnullen 6,024 beträgt, daß also die Wahrscheinlichkeit, eine beliebige Mantisse endige mit Null, auf Grund dieser Erfahrung mit $\frac{6{,}024}{60} = 0{,}1004$ zu bewerten sei, womit empirisch bewiesen ist, daß alle Ziffern (nicht bloß an der End-, sondern an jeder Stelle) gleichberechtigt sind.[1])

187. Beispiel LXVIII. *Verteilung von Brustumfängen.* Unter den Materien, auf die Quetelet seine in Nr. 178 angedeutete Untersuchungsmethode anwandte, befindet sich eine Tabelle der Brustumfänge von 5740 schottischen Soldaten. Sie hat die korrekte Form einer Verteilungstafel und ist in den ersten zwei Kolonnen der folgenden Tabelle mitgeteilt; daran schließt sich die Summenrechnung, die bis zur vierten Ordnung und wieder, mit Rücksicht auf die geringe Ausdehnung der Tafel, einheitlich geführt ist.

x_i engl. Zoll	z_i	$s_i^{(1)}$	$s_i^{(2)}$	$s_i^{(3)}$	$s_i^{(4)}$
33	3	3	3	3	3
34	18	21	24	27	30
35	81	102	126	153	183
36	185	287	413	566	749
37	420	707	1120	1686	2435
38	749	1456	2576	4262	6697
39	1075	2531	5107	9369	16066
40	1079	3610	8717	18086	34152
41	934	4544	13261	31347	65499
42	658	5202	18463	49810	115309
43	370	5572	24035	73845	189154
44	92	5664	29699	103544	430277
45	50	5714	35413	292698	
46	21	5735	138957		
47	4	41148			
48	1				
	5740				

1) Man vergleiche hierzu die Ausführungen H. Poincarés in La Science et l'Hypothèse, deutsch von F. und L. Lindemann, 2. Aufl., Leipzig 1906, S. 193.

Sie liefert die folgenden Grundzahlen:

$$S_0^- = 5735 + 4 = 5739, \quad S_1^- = 41\,148,$$
$$S_2^- = 138\,957, \quad S_3^- = 292\,698, \quad S_4^- = 430277;$$

aus diesen ergibt sich mittels der Gleichungen (40), worin $\delta = 1$ zu setzen ist:

$$m = 5739 + 1 = 5740,$$
$$\eta_1 = -8{,}17$$
$$\eta_2^2 = 70{,}923$$
$$\eta_3^3 = -647{,}639$$
$$\eta_4^4 = 6177{,}585,$$

daraus:

$$A = 48 - 8{,}17 = 39{,}83 \text{ engl. Zoll}^{1)} \; (= 1{,}011 \text{ m})$$

und weiter nach den Formeln (43):

$$\mu_2^2 = 4{,}174, \quad \mu_3^3 = 0{,}006, \quad \mu_4^4 = 50{,}691;$$

sonach ist

$$\text{str}\,(x) = \mu_2 = 2{,}04 \text{ engl. Zoll} \; (= 51{,}8 \text{ mm})$$

$$h = \frac{1}{\text{str}\,(x)\sqrt{2}} = 0{,}346.$$

Hiermit erhält man in weiterer Rechnung nach den Formeln (25):

$$D_3 = -0{,}00017, \quad D_4 = -0{,}00754,$$

so daß die empirische Formel zur Darstellung der Verteilung bei diesem Grade der Approximation lautet:

$$2\mathfrak{S}(X) - 1 = \Phi(V) - 0{,}00017\,\frac{\Phi(V)_3}{2^2} - 0{,}00754\,\frac{\Phi(V)_4}{2^3}.$$

Ihre Prüfung erfolgt am zweckmäßigsten an der Summentafel. Die Summentafel zur vorstehenden Verteilungstafel ergibt sich aus der mit $s_i^{(1)}$ überschriebenen Kolonne durch Division mit $m = 5740$; sie ist in der zweiten Kolonne der nachfolgenden Tabelle enthalten, wobei zu beachten ist, daß nicht die Argumentwerte der Verteilungstafel, sondern die ihnen nachfolgenden Wechselpunkte in Betracht kommen. Für diese Wechselpunkte ist nun $\mathfrak{S}(X)$ nach der vorstehenden Formel zu berechnen. Die bezügliche Rechnung ist nachstehend an zwei Fällen mit allen Details erläutert.

---

1) 1 engl. Zoll = 0,0253998 m.

$$\begin{array}{rl} X = & 36{,}5 \\ A = & 39{,}83 \\ \hline X - A = & -3{,}33 \\ V = h(X-A) = & -1{,}152 \\ \Phi(V) = & -0{,}8967 \\ D_3 \frac{\Phi(V)_3}{4} = & -0{,}00003 \\ D_4 \frac{\Phi(V)_4}{8} = & 0{,}00443 \\ \hline 2\mathfrak{S}(X) - 1 = & -0{,}89230 \\ 2\mathfrak{S}(X) = & 0{,}10770 \\ \mathfrak{S}(X) = & 0{,}054 \end{array} \qquad \begin{array}{rl} X = & 40{,}5 \\ A = & 39{,}83 \\ \hline X - A = & 0{,}67 \\ V = h(X-A) = & 0{,}232 \\ \Phi(V) = & 0{,}2572 \\ D_3 \frac{\Phi(V)_3}{4} = & 0{,}00008 \\ D_4 \frac{\Phi(V)_4}{8} = & -0{,}02704 \\ \hline 2\mathfrak{S}(X) - 1 = & 0{,}23024 \\ 2\mathfrak{S}(X) = & 1{,}23024 \\ \mathfrak{S}(X) = & 0{,}615\,. \end{array}$$

Hierzu sei bemerkt, daß $\Phi(V)$, $\Phi(V)_2$, $\Phi(V)_4, \ldots$ ungerade, $\Phi(V)_1$, $\Phi(V)_3, \ldots$ gerade Funktionen sind, wie man sich bezüglich $\Phi(V)$ an dem definierenden Integral

$$\frac{2}{\sqrt{\pi}} \int_0^V e^{-t^2} dt$$

und bezüglich der andern durch Differentiation überzeugt.

Die folgende Tabelle gibt den Vergleich der berechneten mit den beobachteten Werten von $\mathfrak{S}(X)$.

X	$\mathfrak{S}(X)$ berechnet	$\mathfrak{S}(X)$ beobachtet
35,5	0,012	0,018
36,5	0,054	0,050
37,5	0,142	0,123
38,5	0,278	0,254
39,5	0,448	0,441
40,5	0,615	0,629
41,5	0,773	0,792
42,5	0,896	0,906
43,5	0,965	0,971

188. Beispiel LXIX. *Altersverteilung der Gestorbenen.* Wenn man die Sterbefälle, die in einem bestimmten Personenkreise sich ereignen, nach dem Alter der Gestorbenen verzeichnet, so erhält man einen stetigen Kollektivgegenstand: Exemplare sind die einzelnen Sterbefälle, Argument das Sterbealter.

Die nachstehende Tabelle gibt in den ersten zwei Kolonnen die Verteilung der 62 765 Sterbefälle, die in dem Material, das der Konstruktion der österreichischen Sterblichkeitstafeln[1]) zugrunde lag, unter männlichen versicherten Personen beobachtet worden sind. Das niedrigste beobachtete Sterbealter fiel also in das Intervall von $17\frac{1}{2}$ bis $18\frac{1}{2}$ Jahren, das höchste in das Intervall von $92\frac{1}{2}$ bis $93\frac{1}{2}$ Jahren. Die Beobachtungszahlen z sind nicht das unmittelbare Ergebnis einer Zählung, gingen vielmehr aus einer Aufteilungsrechnung hervor, die in dem Zählmodus ihren Grund hatte; daraus erklären sich die wiederholt auftretenden 0,5.

An die Verteilungstafel schließt sich die Summentafel an, die hier wegen der beträchtlichen Ausdehnung zweiteilig angelegt ist und vom Argumentwert $a = 55$ ausgeht. Sie führt zu folgenden Zahlen:

$$S_0^+ = 30\,024 + 1\,680 = 31\,704 \qquad S_1^+ = 333\,694{,}5$$
$$S_0^- = 27\,761{,}5 + 1\,635{,}5 = 29\,397 \qquad S_1^- = 294\,931$$
$$S_2^+ = 2\,460\,621 \qquad S_3^+ = 13\,944\,656{,}5 \qquad S_4^+ = 63\,985\,502{,}5$$
$$S_2^- = 2\,086\,815{,}5 \qquad S_3^- = 11\,391\,839 \qquad S_4^- = 50\,370\,893{,}5.$$

Daraus leiten sich die zur Durchführung von (40) erforderlichen Zahlen ab:

$$\Sigma_0 = 61\,101 \qquad \Sigma_1 = 628\,625{,}5$$
$$\Delta_0 = 2\,307 \qquad \Delta_1 = 38\,763{,}5$$
$$\Sigma_2 = 4\,547\,436{,}5 \qquad \Sigma_3 = 25\,336\,495{,}5 \qquad \Sigma_4 = 114\,356\,396$$
$$\Delta_2 = 373\,805{,}5 \qquad \Delta_3 = 2\,552\,817{,}5 \qquad \Delta_4 = 13\,614\,609;$$

man findet

$$\begin{aligned} m &= 61\,101 + 1664 = 62\,765 \\ \eta_1 &= 0{,}654 \\ \eta_2^2 &= 175{,}924 \\ \eta_3^3 &= 319{,}863 \\ \eta_4^4 &= 43\,727{,}46 \\ A &= 55 + 0{,}654 = 55{,}654; \end{aligned}$$

letztere Zahl bedeutet also das durchschnittliche Sterbealter des genannten Personenkreises.

Mittels der Formeln (43) erhält man weiter:

$$\mu_2^2 = 175{,}496 \qquad \mu_3^3 = -24{,}926 \qquad \mu_4^4 = 45016{,}09$$

$$\operatorname{str}(x) = \mu_2 = 13{,}25, \quad h = \frac{1}{\operatorname{str}(x)\sqrt{2}} = 0{,}05337.$$

Die sechsfache Streuung beträgt $79\frac{1}{2}$ Jahre, das Intervall zwischen den äußersten Wechselpunkten 76 Jahre.

---

1) Absterbeordnungen aus Beobachtungen an österreichischen Versicherten. 4 Bde., Wien 1907. — Vgl. hierzu die zu Nr. 198 zitierte Abhandlung des Verfassers.

Verteilungstafel		Summentafel			
x	z	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$
18	1	1	1	1	1
19	3	4	5	6	7
20	5,5	9,5	14,5	20,5	27,5
21	9	18,5	33	53,5	81
22	17,5	36	69	122,5	203,5
23	31	67	136	258,5	462
24	55	122	258	516,5	978,5
25	82,5	204,5	462,5	979	1957,5
26	121	325,5	788	1767	3724,5
27	176	501,5	1289,5	3056,5	6781
28	232,5	734	2023,5	5080	11861
29	287	1021	3044,5	8124,5	19985,5
30	336,5	1357,5	4402	12526,5	32512
31	425	1782,5	6184,5	18711	51223
32	502	2284,5	8469	27180	78403
33	576,5	2861	11330	38510	116913
34	671,5	3532,5	14862,5	53372,5	170285,5
35	735,5	4268	19130,5	72503	242788,5
36	803,5	5071,5	24202	96705	339493,5
37	881,5	5953	30155	126860	466353,5
38	940	6893	37048	163908	630261,5
39	1017,5	7910,5	44958,5	208866,5	839128
40	1074,5	8985	53943,5	262810	1101938
41	1128,5	10113,5	64057	326867	1428805
42	1223	11336,5	75393,5	402260,5	1831065,5
43	1283	12619,5	88013	490273,5	2321339
44	1356	13975,5	101988,5	592262	2913601
45	1397	15372,5	117361	709623	3623224
46	1400,5	16773	134134	843757	4466981
47	1502	18275	152409	996166	5463147
48	1567,5	19842,5	172251,5	1168417,5	6631564,5
49	1528	21370,5	193622	1362039,5	7993604
50	1557	22927,5	216549,5	1578589	9572193
51	1580	24507,5	241057	1819646	50370893,5
52	1605	26112,5	267169,5	11391839	
53	1649	27761,5	2086815,5		
54	1635,5	294931			
55	1664				

x	z	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$
56	1680	333694,5			
57	1681,5	30024	2460621		
58	1689,5	28342,5	303670,5	13944656,5	
59	1630,5	26653	275328	2156950,5	63985502,5
60	1562,5	25022,5	248675	1881622,5	11787706
61	1562	23460	223652,5	1632947,5	9906083,5
62	1568	21898	200192,5	1409295	8273136
63	1452,5	20330	178294,5	1209102,5	6863841
64	1404,5	18877,5	157964,5	1030808	5654738,5
65	1466,5	17473	139087	872843,5	4623930,5
66	1465,5	16006,5	121614	733756,5	3751087
67	1386,5	14541	105607,5	612142,5	3017330,5
68	1304,5	13154,5	91066,5	506535	2405188
69	1316	11850	77912	415468,5	1898653
70	1275,5	10534	66062	337556,5	1483184,5
71	1149	9258,5	55528	271494,5	1145628
72	1030,5	8109,5	46269,5	215966,5	874133,5
73	963,5	7079	38160	169697	658167
74	923	6115,5	31081	131537	488470
75	807	5192,5	24965,5	100456	356933
76	732	4385,5	19773	75490,5	256477
77	686	3653,5	15387,5	55717,5	180986,5
78	609	2967,5	11734	40330	125269
79	531,5	2358,5	8766,5	28596	84939
80	423,5	1827	6408	19829,5	56343
81	344,5	1403,5	4581	13421,5	36513,5
82	308,5	1059	3177,5	8840,5	23092
83	249	750,5	2118,5	5663	14251,5
84	176	501,5	1368	3544,5	8588,5
85	126	325,5	866,5	2176,5	5044
86	71	199,5	541	1310	2867,5
87	34,5	128,5	341,5	769	1557,5
88	33,5	94	213	427,5	788,5
89	27,5	60,5	119	214,5	361
90	16,5	33	58,5	95,5	146,5
91	10	16,5	25,5	37	51
92	4	6,5	9	11,5	14
93	2,5	2,5	2,5	2,5	2,5
Verteilungstafel		Summentafel			

Nun sind alle Daten zur Berechnung der D vorhanden; laut (25) ist

$$D_3 = 0{,}00253, \quad D_4 = -0{,}12821;$$

hiermit ergibt sich die empirische Formel zur Darstellung der Verteilung:

$$2\mathfrak{S}(X) - 1 = \Phi(V) + 0{,}00253\,\frac{\Phi(V)_3}{4} - 0{,}12821\,\frac{\Phi(V)_4}{8}.$$

Die Werte von $\mathfrak{S}(X)$, die sie liefert, sind in der nachstehenden, auf Altersquinquennien reduzierten Verteilungstafel mit den aus der Beobachtung hervorgehenden zusammengestellt. Die Wiedergabe ist eine befriedigende; doch deutet der systematische Gang der Abweichungen darauf hin, daß die Formel, wenn es sich um die Erzielung eines vollkommeneren Anschlusses handeln würde, durch Anfügung weiterer Glieder verbessert werden müßte. Man kann immerhin sagen, daß die *prozentische* Aufteilung der Sterbefälle bis zu den einzelnen in der reduzierten Tafel vertretenen Altersgrenzen hin durch die Formel ziemlich gut wiedergegeben wird, wie die folgende Zusammenstellung zeigt:

Altersgrenze:	25,5	30,5	35,5	40,5	45,5	50,5	
Berechnete Prozente:	0,2	2	7	15	25	37	
Beobachtete Prozente:	0,3	2	7	14	24	37	
Altersgrenze:	55,5	60,5	65,5	70,5	75,5	80,5	85,5
Berechnete Prozente:	50	62	74	84	92	97	99,8
Beobachtete Prozente:	49	63	75	85	93	98	99,7.

Reduzierte Verteilungstafel			Summentafel			
Intervall	x	z	Absol. Summen $s^{(1)}$	$\mathfrak{S}(X)$ berechnet	$\mathfrak{S}(X)$ beobachtet	Rechn.-Beobacht.
15,5—20,5	18	9,5	9,5			
20,5—25,5	23	195	204,5	0,0022	0,0033	— 0,0011
25,5—30,5	28	1153	1357,5	0,0243	0,0216	+ 0,0027
30,5—35,5	33	2910,5	4268	0,0723	0,0680	+ 0,0043
35,5—40,5	38	4717	8985	0,1521	0,1432	+ 0,0089
40,5—45,5	43	6387,5	15372,5	0,2526	0,2449	+ 0,0077
45,5—50,5	48	7555	22927,5	0,3743	0,3653	+ 0,0090
50,5—55,5	53	8133,5	31061	0,5035	0,4949	+ 0,0086
55,5—60,5	58	8244	39305	0,6165	0,6262	— 0,0097
60,5—65,5	63	7453,5	46758,5	0,7357	0,7450	— 0,0093
65,5—70,5	68	6748	53506,5	0,8421	0,8525	— 0,0104
70,5—75,5	73	4873	58379,5	0,9237	0,9301	— 0,0064
75,5—80,5	78	2982	61361,5	0,9737	0,9776	— 0,0039
80,5—85,5	83	1204	62565,5	0,9975	0,9968	+ 0,0007
85,5—90,5	88	183	62748,5			
90,5—95,5	93	16,5	62765			
		62765				

189. Die Theoreme von Bernoulli und Poisson unter dem Gesichtspunkte der Kollektivmaßlehre. Es gibt Materien, aus denen sich Kollektivgegenstände mit einer a priori angebbaren theoretischen Verteilungsfunktion konstruieren lassen. Als Typus solcher Materien können planmäßige Ziehungen aus Urnen von bestimmter Füllung mit Kugeln verschiedener Farbe angesehen werden, derart gruppiert, daß sie einen Kollektivgegenstand bilden.

Solche Kollektivgegenstände können zwei prinzipiell verschiedene Verwendungen finden.

Stellt man durch wirkliche Ziehungen, die nach einem Urnenschema ausgeführt worden sind, einen empirischen Kollektivgegenstand her und vergleicht ihn mit dem theoretischen, d. h. seine Verteilungstafel mit der theoretischen Verteilung, so erkennt man das Maß der zufälligen Störungen, durch die sich der empirische Kollektivgegenstand von seinem theoretischen Vorbild entfernt. Auf diesem Grundgedanken beruht die Prüfung wirklicher zufälliger Vorgänge daraufhin, ob sie den Sätzen der Wahrscheinlichkeitsrechnung folgen, in letzter Linie also die induktive Begründung des Gesetzes der großen Zahlen oder des Satzes von der gleichmäßigen Erschöpfung der möglichen Fälle.

Die andere Verwendung besteht darin, daß man einen Kollektivgegenstand beliebiger Provenienz, also irgendeine Massenerscheinung, darauf prüft, ob sie sich unter das Bild eines bestimmten Urnenschemas subsummieren lasse; zu diesem Zwecke formt man aus den Erfahrungen über die Massenerscheinung in passender Weise einen Kollektivgegenstand, legt seine Verteilungstafel an und vergleicht sie mit jener theoretischen Verteilung, zu der das Urnenschema führt. Haben wiederholte Vergleichungen dieser Art eine genügende Übereinstimmung ergeben, so ist damit über die inneren Vorgänge der Massenerscheinung allerdings kein Aufschluß gewonnen, wohl aber eine rein äußerliche, bis zu einem gewissen Grade zutreffende Analogie zwischen der Massenerscheinung und dem Urnenschema nachgewiesen, die sich auf Häufigkeitsfragen und aus diesen gezogene Schlüsse beschränkt. In diesem Sinne angewendet ist die Kollektivmaßlehre ein Mittel, Massenerscheinungen auf die Beständigkeit ihres Verlaufs, ihre *Stabilität*, zu untersuchen und ihre Abhängigkeit von verschiedenen Begleitumständen zu erkennen.

Das Urnenschema, das man früher glaubte auf alle Massenerscheinungen anwenden zu können, ist das Bernoullische, bestehend in wiederholten Ziehungen aus einer Urne mit unverändert bleibender Füllung. Diesem starren Schematismus steht der Poissonsche gegenüber, der wegen der Variabilität der Bedingungen eine größere Anpassungsfähigkeit besitzt.

Bevor wir diese zwei Schemata vom Standpunkte der Kollektiv-

maßlehre betrachten, seien einige Hilfssätze über die $\mathfrak{D}$-Operation vorausgeschickt.

Für zwei unabhängige (unstetige) Argumente x, y seien die Verteilungsfunktionen $\mathfrak{U}_1(x)$, $\mathfrak{U}_2(y)$ gegeben; da diesen die Bedeutung von relativen Häufigkeiten und damit zugleich von Wahrscheinlichkeiten der betreffenden Argumentwerte zukommt, so ist die Verteilungsfunktion $\mathfrak{U}(x, y)$ für die Wertverbindungen von x, y gegeben durch (Nr. 31):

$$\mathfrak{U}(x, y) = \mathfrak{U}_1(x)\mathfrak{U}_2(y).$$

Der mit dieser Verteilung gebildete Durchschnitt einer Funktion $f(x, y)$ derselben Argumente ist

$$\mathfrak{D}[f(x, y)] = \sum\sum f(x, y)\mathfrak{U}(x, y) = \sum\sum f(x, y)\mathfrak{U}_1(x)\mathfrak{U}_2(x),$$

die Summierung über alle Wertverbindungen x, y ausgedehnt.

Ist insbesondere $f(x, y) = \varphi(x) + \psi(y)$, so hat man

$$\mathfrak{D}[\varphi(x) + \psi(y)] = \sum\mathfrak{U}_2(y)\sum\varphi(x)\mathfrak{U}_1(x) + \sum\mathfrak{U}_1(x)\sum\psi(y)\mathfrak{U}_2(y);$$

jedes Summenzeichen erstreckt sich auf alle Werte des ihm nachfolgenden Arguments; da nun $\sum\mathfrak{U}_1(x) = \sum\mathfrak{U}_2(y) = 1$ ist, so ergibt sich die Formel

$$\mathfrak{D}[\varphi(x) + \psi(y)] = \mathfrak{D}[\varphi(x)] + \mathfrak{D}[\psi(y)], \tag{1}$$

die auf eine beliebige Anzahl von Summanden ausgedehnt werden kann.

Ist weiter $f(x, y) = \varphi(x)\psi(y)$, so wird

$$\mathfrak{D}[\varphi(x)\psi(x)] = \sum\varphi(x)\mathfrak{U}_1(x)\sum\psi(y)\mathfrak{U}_2(y) = \mathfrak{D}[\varphi(x)]\mathfrak{D}[\psi(y)], \tag{2}$$

eine Formel, die sich ebenfalls für beliebig viele Faktoren erweisen läßt. Wenn insbesondere $\psi(y) = c$, so wird

$$\mathfrak{D}[c\varphi(x)] = \sum\varphi(x)\mathfrak{U}_1(x)\cdot c\sum\mathfrak{U}_2(y) = c\mathfrak{D}[\varphi(x)]. \tag{3}$$

Für einen stetigen Kollektivgegenstand ergeben sich mittels der Integralrechnung gleichlautende Regeln.

Es seien n Urnen $U_i (i = 1, 2, \cdots n)$ vorhanden, die mit weißen (w) und schwarzen (s) Kugeln derart gefüllt sind, daß

$$\mathfrak{W}_{U_i}(w) = p_i, \; \mathfrak{W}_{U_i}(s) = q_i = 1 - p_i$$

ist. Die einzelne Ziehungsreihe bestehe darin, daß aus jeder Urne eine Kugel hervorgeholt und ihre Farbe verzeichnet wird; solcher Reihen mögen m ausgeführt und nach der relativen Häufigkeit von (w) geordnet werden; es handelt sich um die Verteilungsfunktion dieses Arguments x.

Betrachtet man zuerst die Ziehungen aus einer einzelnen Urne, z. B. U_i, und bezeichnet die relative Häufigkeit von (w) im einzelnen

Zuge mit x_i, so ist dieses Argument nur der zwei Werte 0 und 1 fähig und durchläuft während der Ziehungen eine m-gliedrige, aus 0 und 1 irgendwie zusammengesetzte Reihe; die Verteilung dieser zwei Argumentwerte ist aber theoretisch a priori angebbar, indem

$$\mathfrak{U}_i(0) = \mathfrak{W}_{U_i}(s) = q_i, \quad \mathfrak{U}_i(1) = \mathfrak{W}_{U_i}(w) = p_i;$$

auf Grund dieser bekannten Verteilung findet man

$$\mathfrak{D}(x_i) = 0\,\mathfrak{U}_i(0) + 1\,\mathfrak{U}_i(1) = p_i \tag{4}$$

$$\begin{aligned} str^2(x_i) = \mathfrak{D}[(x_i - p_i)^2] &= (0 - p_i)^2\mathfrak{U}_i(0) + (1 - p_i)^2\mathfrak{U}_i(1) \\ &= p_i^2 q_i + p_i q_i^2 = p_i q_i. \end{aligned} \tag{5}$$

Nun wenden wir uns den Zügen einer Ziehungs*reihe* zu und ordnen wieder jedem die relative Häufigkeit von (w) zu; dann werden die den einzelnen Zügen zugeordneten Zahlen

$$x_1, x_2, \cdots x_i, \cdots x_n$$

wieder teils 0, teils 1 sein, ihre Summe $\sum_1^n x_i$ aber bedeutet die absolute, $\frac{1}{n}\sum_1^n x_i = x$ also die relative Häufigkeit von (w) in der Ziehungsreihe, so daß

$$nx = x_1 + x_2 + \cdots + x_i + \cdots + x_n. \tag{6}$$

Die Verteilung eines jeden einzelnen x_i der rechten Seite ist durch die Elemente: Argumentdurchschnitt und Streuung im Sinne der Gleichungen (4), (5) gekennzeichnet; daraus lassen sich unter Anwendung der vorausgeschickten Regeln (1), (2), (3) die analogen Elemente für die Verteilung von x ableiten. Es ist nämlich

$$\mathfrak{D}(nx) = n\mathfrak{D}(x) = \mathfrak{D}(x_1) + \mathfrak{D}(x_2) + \cdots + \mathfrak{D}(x_n) = p_1 + p_2 + \cdots + p_n,$$

woraus

$$\mathfrak{D}(x) = \frac{p_1 + p_2 + \cdots + p_n}{n} = p; \tag{7}$$

das Komplement hierzu heiße q, dann ist $q = 1 - p = \frac{q_1 + q_2 + \cdots + q_n}{n}$.

Weiter kann mit Rücksicht hierauf aus (6)

$$n(x - p) = (x_1 - p_1) + (x_2 - p_2) + \cdots + (x_n - p_n)$$

gefolgert werden, woraus sich

$$n^2(x - p)^2 = \sum_1^n (x_i - p_i)^2 + \sum (x_i - p_i)(x_j - p_j)$$

ergibt, die zweite Summierung über alle Wertpaare $i, j\,(i \neq j)$ erstreckt; mithin hat man

$$\mathfrak{D}[n^2(x-p)^2] = n^2\mathfrak{D}[(x-p)^2] = \sum_1^n \mathfrak{D}[(x_i-p_i)^2] + \sum \mathfrak{D}(x_i-p_i)\mathfrak{D}(x_j-p_j),$$

und da laut (4) jedes $\mathfrak{D}(x_i - p_i) = 0$ ist, weiter

$$n^2 \operatorname{str}^2(x) = \sum_1^n \operatorname{str}^2(x_i) = \sum_1^n p_i q_i. \tag{8}$$

Setzt man $p_i - p = \delta_i$, so ist $q_i - q = -\delta_i$ und

$$\sum_1^n p_i q_i = \sum_1^n (p + \delta_i)(q - \delta_i) = npq - (p - q)\sum_1^n \delta_i - \sum_1^n \delta_i^2,$$

wegen $\sum \delta_i = 0$ aber reduziert sich dies auf

$$n \operatorname{str}^2(x) = pq - \frac{\sum_1^n \delta_i^2}{n}. \tag{8*}$$

Geht man nun daran, die Verteilung des Arguments x, das zwar unstetig und lediglich der Werte

$$0,\ \frac{1}{n},\ \frac{2}{n},\ \ldots\ \frac{n-1}{n},\ 1$$

fähig ist, bei großem n wegen der dichten Anordnung dieser Werte aber als stetig behandelt werden kann, durch die Φ-Reihe darzustellen, und beschränkt man sich bei der Summenfunktion auf das erste Glied, so daß in erster Näherung

$$2\mathfrak{S}(X) - 1 = \Phi(V)$$

genommen wird, so hat man

$$\mathfrak{V}(X)\,dX = \frac{1}{2}\Phi(V)_1\,dV,$$

wobei

$$V = h(X - \mathfrak{D}(x)) = h(X - p), \quad dV = h\,dX$$

zu setzen ist; hiermit wird

$$\mathfrak{V}(X) = \frac{h}{\sqrt{\pi}} e^{-h^2(X-p)^2}$$

und die Elemente dieser Funktion sind laut (7) und (8):

$$p = \frac{p_1 + p_2 + \cdots + p_n}{n}$$

$$h = \frac{1}{\operatorname{str}(x)\sqrt{2}} = \frac{n}{\sqrt{2\sum_1^n p_i q_i}}.$$

Die Wahrscheinlichkeit, daß die Abweichung $X - p = \theta$ zwischen den Grenzen

$$\mp \frac{\gamma}{h} = \mp \gamma \frac{\sqrt{2 \sum_{1}^{n} p_i q_i}}{n}$$

eingeschlossen sei, beträgt demnach

$$\frac{h}{\sqrt{\pi}} \int_{-\frac{\gamma}{h}}^{\frac{\gamma}{h}} e^{-h^2 \theta^2} d\theta = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-t^2} dt.$$

Hiermit ist der hauptsächliche Inhalt des *Poissonschen Theorems* wieder gefunden; die Deduktion hat vor der Poissonschen (Nr. 93) den Vorzug, daß sie sich an eine allgemeinere Gedankenbildung anlehnt und das Theorem als einen speziellen Fall erkennen läßt, der sich aus einer umfassenderen Fragestellung ergibt. Auch daß man es bei der vorstehenden Formulierung mit einer ersten Näherung zu tun hat, die aber schon bei einigermaßen großer Versuchszahl n für praktische Bedürfnisse ausreicht, ist aus dem Gange der Entwicklung zu ersehen.[1])

Der Übergang zum *Bernoullischen Theorem* ist unmittelbar vollzogen; man braucht nur die Urnen als von gleicher Füllung oder alle Ziehungen aus *einer* Urne ausgeführt vorauszusetzen, deren $\mathfrak{W}(w) = p$, $\mathfrak{W}(s) = q = 1 - p$ sei; dann ergibt sich für die Grenzen

$$\mp \frac{\gamma}{h} = \mp \gamma \sqrt{\frac{2pq}{n}}$$

der Abweichung vom normalen Wert p die Wahrscheinlichkeit

$$\frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-t^2} dt$$

in Übereinstimmung mit dem wesentlichen Inhalt des Bernoullischen Theorems (Nr. 77).

Bemerkenswert ist die Streuungsformel (8*); sie zeigt die schon an früherer Stelle (Nr. 95) in anderem Zusammenhange festgestellte Tatsache, daß bei dem Poissonschen Urnenschema die Streuung kleiner ist als bei einem Bernoullischen mit denselben Grundwahr-

1) Bruns, a. a. O. S. 193—199, hat auch die Fortsetzung der Φ-Reihe aus den Daten des Urnenschemas a priori festgestellt, von der bei strengeren Anforderungen Gebrauch zu machen wäre.

scheinlichkeiten p, q, wo $n \operatorname{str}^2(x) = pq$ ist, und zwar um so kleiner, je größer das Glied

$$\frac{\sum_1^n \delta_i^2}{n}$$

ausfällt; dieses Glied stellt aber vermöge der Bedeutung des δ_i das Quadrat der mittleren Abweichung der p_i von dem normalen Wert p, also die Streuung der p_i. Je verschiedener also das Füllungsverhältnis der Urnen, um so kleiner wird die Streuung (8*). Bei dem gleichen p kann sich die Streuung sehr verschieden gestalten je nach der Verteilung der p_i; darin liegt der Grund für die Anpassungsfähigkeit des Poissonschen Schemas an verschiedene Verteilungen.

Tafel I.

Werte der Funktion $\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$

x	$\Phi(x)$	Diff.	x	$\Phi(x)$	Diff.	x	$\Phi(x)$	Diff.
0,00	0,0000 000	112 833	0,40	0,4283 922	95 768	0,80	0,7421 010	59 023
0,01	0,0112 833	112 811	0,41	0,4379 690	94 986	0,81	0,7480 033	58 075
0,02	0,0225 644	112 766	0,42	0,4474 676	94 191	0,82	0,7538 108	57 130
0,03	0,0338 410	112 699	0,43	0,4568 867	93 384	0,83	0,7595 238	56 189
0,04	0,0451 109	112 609	0,44	0,4662 251	92 567	0,84	0,7651 427	55 253
0,05	0,0563 718	112 497	0,45	0,4754 818	91 737	0,85	0,7706 680	54 322
0,06	0,0676 215	112 362	0,46	0,4846 555	90 897	0,86	0,7761 022	53 396
0,07	0,0788 577	112 204	0,47	0,4937 452	90 046	0,87	0,7814 398	52 475
0,08	0,0900 781	112 025	0,48	0,5027 498	89 185	0,88	0,7866 873	51 559
0,09	0,1012 806	111 824	0,49	0,5116 683	88 316	0,89	0,7918 432	50 650
0,10	0,1124 630	111 600	0,50	0,5204 999	87 438	0,90	0,7969 082	49 746
0,11	0,1236 230	111 354	0,51	0,5292 437	86 550	0,91	0,8018 828	48 849
0,12	0,1347 584	111 087	0,52	0,5378 987	85 654	0,92	0,8067 677	47 958
0,13	0,1458 671	110 799	0,53	0,5464 641	84 751	0,93	0,8115 635	47 075
0,14	0,1569 470	110 489	0,54	0,5549 392	83 841	0,94	0,8162 710	46 198
0,15	0,1679 959	110 158	0,55	0,5633 233	82 924	0,95	0,8208 908	45 328
0,16	0,1790 117	109 806	0,56	0,5716 157	82 001	0,96	0,8254 236	44 467
0,17	0,1899 923	109 434	0,57	0,5798 158	81 071	0,97	0,8298 703	43 612
0,18	0,2009 357	109 041	0,58	0,5879 229	80 136	0,98	0,8342 315	42 766
0,19	0,2118 398	108 627	0,59	0,5959 365	79 196	0,99	0,8385 081	41 927
0,20	0,2227 025	108 193	0,60	0,6038 561	78 251	1,00	0,8427 008	41 097
0,21	0,2335 218	107 740	0,61	0,6116 812	77 302	1,01	0,8468 105	40 275
0,22	0,2442 958	107 267	0,62	0,6194 114	76 349	1,02	0,8508 380	39 462
0,23	0,2550 225	106 775	0,63	0,6270 463	75 394	1,03	0,8547 842	38 657
0,24	0,2657 000	106 263	0,64	0,6345 857	74 435	1,04	0,8586 499	37 861
0,25	0,2763 263	105 734	0,65	0,6420 292	73 473	1,05	0,8624 360	37 075
0,26	0,2868 997	105 185	0,66	0,6493 765	72 510	1,06	0,8661 435	36 297
0,27	0,2974 182	104 618	0,67	0,6566 275	71 545	1,07	0,8697 732	35 529
0,28	0,3078 800	104 034	0,68	0,6637 820	70 579	1,08	0,8733 261	34 769
0,29	0,3182 834	103 433	0,69	0,6708 399	69 611	1,09	0,8768 030	34 020
0,30	0,3286 267	102 814	0,70	0,6778 010	68 644	1,10	0,8802 050	33 280
0,31	0,3389 081	102 178	0,71	0,6846 654	67 676	1,11	0,8835 330	32 549
0,32	0,3491 259	101 526	0,72	0,6914 330	66 708	1,12	0,8867 879	31 828
0,33	0,3592 785	100 859	0,73	0,6981 038	65 742	1,13	0,8899 707	31 116
0,34	0,3693 644	100 175	0,74	0,7046 780	64 776	1,14	0,8930 823	30 415
0,35	0,3793 819	99 477	0,75	0,7111 556	63 811	1,15	0,8961 238	29 724
0,36	0,3893 296	98 763	0,76	0,7175 367	62 849	1,16	0,8990 962	29 042
0,37	0,3992 059	98 034	0,77	0,7238 216	61 888	1,17	0,9020 004	28 370
0,38	0,4090 093	97 292	0,78	0,7300 104	60 931	1,18	0,9048 374	27 709
0,39	0,4187 385	96 537	0,79	0,7361 035	59 975	1,19	0,9076 083	27 057

x	$\Phi(x)$	Diff.	x	$\Phi(x)$	Diff.	x	$\Phi(x)$	Diff.
1,20	0,9103 140	26 415	1,70	0,9837 904	6 166	2,20	0,9981 372	872
1,21	0,9129 555	25 784	1,71	0,9844 070	5 958	2,21	0,9982 244	835
1,22	0,9155 339	25 162	1,72	0,9850 028	5 757	2,22	0,9983 079	799
1,23	0,9180 501	24 551	1,73	0,9855 785	5 561	2,23	0,9983 878	764
1,24	0,9205 052	23 949	1,74	0,9861 346	5 371	2,24	0,9984 642	731
1,25	0,9229 001	23 358	1,75	0,9866 717	5 186	2,25	0,9985 373	698
1,26	0,9252 359	22 777	1,76	0,9871 903	5 007	2,26	0,9986 071	668
1,27	0,9275 136	22 206	1,77	0,9876 910	4 832	2,27	0,9986 739	638
1,28	0,9297 342	21 645	1,78	0,9881 742	4 664	2,28	0,9987 377	609
1,29	0,9318 987	21 093	1,79	0,9886 406	4 499	2,29	0,9987 986	582
1,30	0,9340 080	20 552	1,80	0,9890 905	4 340	2,30	0,9988 568	556
1,31	0,9360 632	20 020	1,81	0,9895 245	4 186	2,31	0,9989 124	531
1,32	0,9380 652	19 498	1,82	0,9899 431	4 036	2,32	0,9989 655	507
1,33	0,9400 150	18 987	1,83	0,9903 467	3 892	2,33	0,9990 162	484
1,34	0,9419 137	18 485	1,84	0,9907 359	3 751	2,34	0,9990 646	461
1,35	0,9437 622	17 992	1,85	0,9911 110	3 615	2,35	0,9991 107	441
1,36	0,9455 614	17 510	1,86	0,9914 725	3 482	2,36	0,9991 548	420
1,37	0,9473 124	17 036	1,87	0,9918 207	3 355	2,37	0,9991 968	401
1,38	0,9490 160	16 573	1,88	0,9921 562	3 231	2,38	0,9992 369	382
1,39	0,9506 733	16 118	1,89	0,9924 793	3 111	2,39	0,9992 751	364
1,40	0,9522 851	15 673	1,90	0,9927 904	2 995	2,40	0,9993 115	347
1,41	0,9538 524	15 238	1,91	0,9930 899	2 883	2,41	0,9993 462	331
1,42	0,9553 762	14 811	1,92	0,9933 782	2 775	2,42	0,9993 793	315
1,43	0,9568 573	14 393	1,93	0,9936 557	2 664	2,43	0,9994 108	300
1,44	0,9582 966	13 984	1,94	0,9939 226	2 568	2,44	0,9994 408	286
1,45	0,9596 950	13 585	1,95	0,9941 794	2 469	2,45	0,9994 694	272
1,46	0,9610 535	13 194	1,96	0,9944 263	2 374	2,46	0,9994 966	260
1,47	0,9623 729	12 812	1,97	0,9946 637	2 283	2,47	0,9995 226	246
1,48	0,9636 541	12 483	1,98	0,9948 920	2 194	2,48	0,9995 472	235
1,49	0,9648 979	12 073	1,99	0,9951 114	2 109	2,49	0,9995 707	223
1,50	0,9661 052	11 716	2,00	0,9953 223	2 025	2,50	0,9995 930	213
1,51	0,9672 768	11 367	2,01	0,9955 248	1 947	2,51	0,9996 143	202
1,52	0,9684 135	11 027	2,02	0,9957 195	1 868	2,52	0,9996 345	192
1,53	0,9695 162	10 695	2,03	0,9959 063	1 795	2,53	0,9996 537	183
1,54	0,9705 857	10 370	2,04	0,9960 858	1 733	2,54	0,9996 720	173
1,55	0,9716 227	10 054	2,05	0,9962 581	1 654	2,55	0,9996 893	165
1,56	0,9726 281	9 745	2,06	0,9964 235	1 587	2,56	0,9997 058	157
1,57	0,9736 026	9 444	2,07	0,9965 822	1 522	2,57	0,9997 215	149
1,58	0,9745 470	9 150	2,08	0,9967 344	1 461	2,58	0,9997 364	141
1,59	0,9754 620	8 864	2,09	0,9968 805	1 400	2,59	0,9997 505	135
1,60	0,9763 484	8 585	2,10	0,9970 205	1 343	2,60	0,9997 640	127
1,61	0,9772 069	8 312	2,11	0,9971 548	1 288	2,61	0,9997 767	121
1,62	0,9780 381	8 048	2,12	0,9972 836	1 234	2,62	0,9997 888	115
1,63	0,9788 429	7 789	2,13	0,9974 070	1 183	2,63	0,9998 003	109
1,64	0,9796 218	7 538	2,14	0,9975 253	1 133	2,64	0,9998 112	103
1,65	0,9803 756	7 293	2,15	0,9976 386	1 086	2,65	0,9998 215	98
1,66	0,9811 049	7 055	2,16	0,9977 472	1 039	2,66	0,9998 313	93
1,67	0,9818 104	6 824	2,17	0,9978 511	994	2,67	0,9998 406	88
1,68	0,9824 928	6 598	2,18	0,9979 505	954	2,68	0,9998 494	84
1,69	0,9831 526	6 378	2,19	0,9980 459	913	2,69	0,9998 578	79

x	$\Phi(x)$	Diff.	x	$\Phi(x)$	Diff.	x	$\Phi(x)$
2,70	0,9998 657	75	3,20	0,9999 940	4	3,70	0,9999 9983 285
2,71	0,9998 732	71	3,21	0,9999 944	3	3,71	0,9999 9984 517
2,72	0,9998 802	67	3,22	0,9999 947	4	3,72	0,9999 9985 663
2,73	0,9998 870	63	3,23	0,9999 951	3	3,73	0,9999 9986 726
2,74	0,9998 933	61	3,24	0,9999 954	3	3,74	0,9999 9987 712
2,75	0,9998 994	57	3,25	0,9999 957	3	3,75	0,9999 9988 629
2,76	0,9999 051	54	3,26	0,9999 960	2	3,76	0,9999 9989 477
2,77	0,9999 105	51	3,27	0,9999 962	3	3,77	0,9999 9990 265
2,78	0,9999 156	48	3,28	0,9999 965	2	3,78	0,9999 9990 995
2,79	0,9999 204	46	3,29	0,9999 967	2	3,79	0,9999 9991 672
2,80	0,9999 250	43	3,30	0,9999 969	2	3,80	0,9999 9992 300
2,81	0,9999 293	41	3,31	0,9999 971	2	3,81	0,9999 9992 881
2,82	0,9999 334	38	3,32	0,9999 973	2	3,82	0,9999 9993 421
2,83	0,9999 372	37	3,33	0,9999 975	2	3,83	0,9999 9993 921
2,84	0,9999 409	34	3,34	0,9999 977	1	3,84	0,9999 9994 383
2,85	0,9999 443	33	3,35	0,9999 978	2	3,85	0,9999 9994 812
2,86	0,9999 476	31	3,36	0,9999 980	1	3,86	0,9999 9995 208
2,87	0,9999 507	29	3,37	0,9999 981	1	3,87	0,9999 9995 575
2,88	0,9999 536	27	3,38	0,9999 982	2	3,88	0,9999 9995 915
2,89	0,9999 563	26	3,39	0,9999 984	1	3,89	0,9999 9996 230
2,90	0,9999 589	24	3,40	0,9999 985	1	3,90	0,9999 9996 521
2,91	0,9999 613	23	3,41	0,9999 986	1	3,91	0,9999 9996 790
2,92	0,9999 636	22	3,42	0,9999 987	1	3,92	0,9999 9997 039
2,93	0,9999 658	21	3,43	0,9999 988	1	3,93	0,9999 9997 260
2,94	0,9999 679	19	3,44	0,9999 989	0	3,94	0,9999 9997 482
2,95	0,9999 698	18	3,45	0,9999 989		3,95	0,9999 9997 678
2,96	0,9999 716	17	3,46	0,9999 9900 780		3,96	0,9999 9997 860
2,97	0,9999 733	17	3,47	0,9999 9907 672		3,97	0,9999 9998 028
2,98	0,9999 750	15	3,48	0,9999 9914 101		3,98	0,9999 9998 183
2,99	0,9999 765	14	3,49	0,9999 9920 097		3,99	0,9999 9998 327
3,00	0,9999 779	14	3,50	0,9999 9925 691		4,0	0,9999 9998 458
3,01	0,9999 793	12	3,51	0,9999 9930 905		4,1	0,9999 9999 330
3,02	0,9999 805	12	3,52	0,9999 9935 766		4,2	0,9999 9999 714
3,03	0,9999 817	12	3,53	0,9999 9940 296		4,3	0,9999 9999 881
3,04	0,9999 829	10	3,54	0,9999 9944 519		4,4	0,9999 9999 951
3,05	0,9999 839	10	3,55	0,9999 9948 452		4,5	0,9999 9999 980
3,06	0,9999 849	10	3,56	0,9999 9952 115		4,6	0,9999 9999 992
3,07	0,9999 859	8	3,57	0,9999 9955 527		4,7	0,9999 9999 997
3,08	0,9999 867	9	3,58	0,9999 9958 703		4,8	0,9999 9999 999
3,09	0,9999 876	8	3,59	0,9999 9961 661			
3,10	0,9999 884	7	3,60	0,9999 9964 414			
3,11	0,9999 891	7	3,61	0,9999 9966 975			
3,12	0,9999 898	6	3,62	0,9999 9969 358			
3,13	0,9999 904	6	3,63	0,9999 9971 574			
3,14	0,9999 910	6	3,64	0,9999 9973 636			
3,15	0,9999 916	5	3,65	0,9999 9975 551			
3,16	0,9999 921	5	3,66	0,9999 9977 333			
3,17	0,9999 926	5	3,67	0,9999 9978 990			
3,18	0,9999 931	5	3,68	0,9999 9980 528			
3,19	0,9999 936	4	3,69	0,9999 9981 957			

x	0	1	2	3	4	5	6	7	8	9
0.00	0.0000	0011	0023	0034	0045	0056	0068	0079	0090	0102
01	0.0113	0124	0135	0147	0158	0169	0181	0192	0203	0214
02	0.0226	0237	0248	0259	0271	0282	0293	0305	0316	0327
03	0.0338	0350	0361	0372	0384	0395	0406	0417	0429	0440
04	0.0451	0462	0474	0485	0496	0507	0519	0530	0541	0552
05	0.0564	0575	0586	0597	0609	0620	0631	0642	0654	0665
06	0.0676	0687	0699	0710	0721	0732	0744	0755	0766	0777
07	0.0789	0800	0811	0822	0833	0845	0856	0867	0878	0890
08	0.0901	0912	0923	0934	0946	0957	0968	0979	0990	1002
09	0.1013	1024	1035	1046	1058	1069	1080	1091	1102	1113
0.10	0.1125	1136	1147	1158	1169	1180	1192	1203	1214	1225
11	0.1236	1247	1259	1270	1281	1292	1303	1314	1325	1336
12	0.1348	1359	1370	1381	1392	1403	1414	1425	1436	1448
13	0.1459	1470	1481	1492	1503	1514	1525	1536	1547	1558
14	0.1569	1581	1592	1603	1614	1625	1636	1647	1658	1669
15	0.1680	1691	1702	1713	1724	1735	1746	1757	1768	1779
16	0.1790	1801	1812	1823	1834	1845	1856	1867	1878	1889
17	0.1900	1911	1922	1933	1944	1955	1966	1977	1988	1998
18	0.2009	2020	2031	2042	2053	2064	2075	2086	2097	2108
19	0.2118	2129	2140	2151	2162	2173	2184	2194	2205	2216
0.20	0.2227	2238	2249	2260	2270	2281	2292	2303	2314	2324
21	0.2335	2346	2357	2368	2378	2389	2400	2411	2421	2432
22	0.2443	2454	2464	2475	2486	2497	2507	2518	2529	2540
23	0.2550	2561	2572	2582	2593	2604	2614	2625	2636	2646
24	0.2657	2668	2678	2689	2700	2710	2721	2731	2742	2753
25	0.2763	2774	2784	2795	2806	2816	2827	2837	2848	2858
26	0.2869	2880	2890	2901	2911	2922	2932	2943	2953	2964
27	0.2974	2985	2995	3006	3016	3027	3037	3047	3058	3068
28	0.3079	3089	3100	3110	3120	3131	3141	3152	3162	3172
29	0.3183	3193	3204	3214	3224	3235	3245	3255	3266	3276
0.30	0.3286	3297	3307	3317	3327	3338	3348	3358	3369	3379
31	0.3389	3399	3410	3420	3430	3440	3450	3461	3471	3481
32	0.3491	3501	3512	3522	3532	3542	3552	3562	3573	3583
33	0.3593	3603	3613	3623	3633	3643	3653	3663	3674	3684
34	0.3694	3704	3714	3724	3734	3744	3754	3764	3774	3784
35	0.3794	3804	3814	3824	3834	3844	3854	3864	3873	3883
36	0.3893	3903	3913	3923	3933	3943	3953	3963	3972	3982
37	0.3992	4002	4012	4022	4031	4041	4051	4061	4071	4080
38	0.4090	4100	4110	4119	4129	4139	4149	4158	4168	4178
39	0.4187	4197	4207	4216	4226	4236	4245	4255	4265	4274
0.40	0.4284	4294	4303	4313	4322	4332	4341	4351	4361	4370
41	0.4380	4389	4399	4408	4418	4427	4437	4446	4456	4465
42	0.4475	4484	4494	4503	4512	4522	4531	4541	4550	4559
43	0.4569	4578	4588	4597	4606	4616	4625	4634	4644	4653
44	0.4662	4672	4681	4690	4699	4709	4718	4727	4736	4746
45	0.4755	4764	4773	4782	4792	4801	4810	4819	4828	4837
46	0.4847	4856	4865	4874	4883	4892	4901	4910	4919	4928
47	0.4937	4946	4956	4965	4974	4983	4992	5001	5010	5019
48	0.5027	5036	5045	5054	5063	5072	5081	5090	5099	5108
49	0.5117	5126	5134	5143	5152	5161	5170	5179	5187	5196
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
51	0.5292	5301	5310	5318	5327	5336	5344	5353	5362	5370
52	0.5379	5388	5396	5405	5413	5422	5430	5439	5448	5456
53	0.5465	5473	5482	5490	5499	5507	5516	5524	5533	5541
54	0.5549	5558	5566	5575	5583	5591	5600	5608	5617	5625
55	0.5633	5642	5650	5658	5667	5675	5683	5691	5700	5708
56	0.5716	5724	5733	5741	5749	5757	5765	5774	5782	5790
57	0.5798	5806	5814	5823	5831	5839	5847	5855	5863	5871
58	0.5879	5887	5895	5903	5911	5919	5927	5935	5943	5951
59	0.5959	5967	5975	5983	5991	5999	6007	6015	6023	6031
0.60	0.6039	6046	6054	6062	6070	6078	6086	6093	6101	6109
61	0.6117	6125	6132	6140	6148	6156	6163	6171	6179	6186
62	0.6194	6202	6209	6217	6225	6232	6240	6248	6255	6263
63	0.6270	6278	6286	6293	6301	6308	6316	6323	6331	6338
64	0.6346	6353	6361	6368	6376	6383	6391	6398	6405	6413
65	0.6420	6428	6435	6442	6450	6457	6464	6472	6479	6486
66	0.6494	6501	6508	6516	6523	6530	6537	6545	6552	6559
67	0.6566	6573	6581	6588	6595	6602	6609	6616	6624	6631
68	0.6638	6645	6652	6659	6666	6673	6680	6687	6694	6701
69	0.6708	6715	6722	6729	6736	6743	6750	6757	6764	6771
0.70	0.6778	6785	6792	6799	6806	6812	6819	6826	6833	6840
71	0.6847	6853	6860	6867	6874	6881	6887	6894	6901	6908
72	0.6914	6921	6928	6934	6941	6948	6954	6961	6968	6974
73	0.6981	6988	6994	7001	7007	7014	7021	7027	7034	7040
74	0.7047	7053	7060	7066	7073	7079	7086	7092	7099	7105
75	0.7112	7118	7124	7131	7137	7144	7150	7156	7163	7169
76	0.7175	7182	7188	7194	7201	7207	7213	7219	7226	7232
77	0.7238	7244	7251	7257	7263	7269	7275	7282	7288	7294
78	0.7300	7306	7312	7318	7325	7331	7337	7343	7349	7355
79	0.7361	7367	7373	7379	7385	7391	7397	7403	7409	7415
0.80	0.7421	7427	7433	7439	7445	7451	7457	7462	7468	7474
81	0.7480	7486	7492	7498	7503	7509	7515	7521	7527	7532
82	0.7538	7544	7550	7555	7561	7567	7572	7578	7584	7590
83	0.7595	7601	7607	7612	7618	7623	7629	7635	7640	7646
84	0.7651	7657	7663	7668	7674	7679	7685	7690	7696	7701
85	0.7707	7712	7718	7723	7729	7734	7739	7745	7750	7756
86	0.7761	7766	7772	7777	7782	7788	7793	7798	7804	7809
87	0.7814	7820	7825	7830	7835	7841	7846	7851	7856	7862
88	0.7867	7872	7877	7882	7888	7893	7898	7903	7908	7913
89	0.7918	7924	7929	7934	7939	7944	7949	7954	7959	7964
0.90	0.7969	7974	7979	7984	7989	7994	7999	8004	8009	8014
91	0.8019	8024	8029	8034	8038	8043	8048	8053	8058	8063
92	0.8068	8073	8077	8082	8087	8092	8097	8101	8106	8111
93	0.8116	8120	8125	8130	8135	8139	8144	8149	8153	8158
94	0.8163	8167	8172	8177	8181	8186	8191	8195	8200	8204
95	0.8209	8213	8218	8223	8227	8232	8236	8241	8245	8250
96	0.8254	8259	8263	8268	8272	8277	8281	8285	8290	8294
97	0.8299	8303	8307	8312	8316	8321	8325	8329	8334	8338
98	0.8342	8347	8351	8355	8360	8364	8368	8372	8377	8381
99	0.8385	8389	8394	8398	8402	8406	8410	8415	8419	8423
1.00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
1.00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
01	0.8468	8472	8476	8480	8484	8488	8492	8496	8500	8504
02	0.8508	8512	8516	8520	8524	8528	8532	8536	8540	8544
03	0.8548	8552	8556	8560	8563	8567	8571	8575	8579	8583
04	0.8586	8590	8594	8598	8602	8606	8609	8613	8617	8621
05	0.8624	8628	8632	8636	8639	8643	8647	8650	8654	8658
06	0.8661	8665	8669	8672	8676	8680	8683	8687	8691	8694
07	0.8698	8701	8705	8708	8712	8716	8719	8723	8726	8730
08	0.8733	8737	8740	8744	8747	8751	8754	8758	8761	8765
09	0.8768	8771	8775	8778	8782	8785	8789	8792	8795	8799
1.10	0.8802	8805	8809	8812	8815	8819	8822	8825	8829	8832
11	0.8835	8839	8842	8845	8848	8852	8855	8858	8861	8865
12	0.8868	8871	8874	8878	8881	8884	8887	8890	8893	8897
13	0.8900	8903	8906	8909	8912	8915	8918	8922	8925	8928
14	0.8931	8934	8937	8940	8943	8946	8949	8952	8955	8958
15	0.8961	8964	8967	8970	8973	8976	8979	8982	8985	8988
16	0.8991	8994	8997	9000	9003	9006	9008	9011	9014	9017
17	0.9020	9023	9026	9029	9031	9034	9037	9040	9043	9046
18	0.9048	9051	9054	9057	9060	9062	9065	9068	9071	9073
19	0.9076	9079	9082	9084	9087	9090	9092	9095	9098	9100
1.20	0.9103	9106	9108	9111	9114	9116	9119	9122	9124	9127
21	0.9130	9132	9135	9137	9140	9143	9145	9148	9150	9153
22	0.9155	9158	9160	9163	9165	9168	9171	9173	9176	9178
23	0.9181	9183	9185	9188	9190	9193	9195	9198	9200	9203
24	0.9205	9207	9210	9212	9215	9217	9219	9222	9224	9227
25	0.9229	9231	9234	9236	9238	9241	9243	9245	9248	9250
26	0.9252	9255	9257	9259	9262	9264	9266	9268	9271	9273
27	0.9275	9277	9280	9282	9284	9286	9289	9291	9293	9295
28	0.9297	9300	9302	9304	9306	9308	9310	9313	9315	9317
29	0.9319	9321	9323	9325	9327	9330	9332	9334	9336	9338
1.30	0.9340	9342	9344	9346	9348	9350	9352	9355	9357	9359
31	0.9361	9363	9365	9367	9369	9371	9373	9375	9377	9379
32	0.9381	9383	9385	9387	9389	9390	9392	9394	9396	9398
33	0.9400	9402	9404	9406	9408	9410	9412	9413	9415	9417
34	0.9419	9421	9423	9425	9427	9428	9430	9432	9434	9436
35	0.9438	9439	9441	9443	9445	9447	9448	9450	9452	9454
36	0.9456	9457	9459	9461	9463	9464	9466	9468	9470	9471
37	0.9473	9475	9477	9478	9480	9482	9483	9485	9487	9488
38	0.9490	9492	9494	9495	9497	9499	9500	9502	9503	9505
39	0.9507	9508	9510	9512	9513	9515	9516	9518	9520	9521
1.40	0.9523	9524	9526	9528	9529	9531	9532	9534	9535	9537
41	0.9539	9540	9542	9543	9545	9546	9548	9549	9551	9552
42	0.9554	9555	9557	9558	9560	9561	9563	9564	9566	9567
43	0.9569	9570	9571	9573	9574	9576	9577	9579	9580	9582
44	0.9583	9584	9586	9587	9589	9590	9591	9593	9594	9596
45	0.9597	9598	9600	9601	9602	9604	9605	9607	9608	9609
46	0.9611	9612	9613	9615	9616	9617	9618	9620	9621	9622
47	0.9624	9625	9626	9628	9629	9630	9631	9633	9634	9635
48	0.9637	9638	9639	9640	9642	9643	9644	9645	9647	9648
49	0.9649	9650	9651	9653	9654	9655	9656	9657	9659	9660
1.50	0.9661	9662	9663	9665	9666	9667	9668	9669	9670	9672
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
1.5	0.9661	9673	9684	9695	9706	9716	9726	9736	9745	9755
1.6	0.9763	9772	9780	9788	9796	9804	9811	9818	9825	9832
1.7	0.9838	9844	9850	9856	9861	9867	9872	9877	9882	9886
1.8	0.9891	9895	9899	9903	9907	9911	9915	9918	9922	9925
1.9	0.9928	9931	9934	9937	9939	9942	9944	9947	9949	9951
2.0	0.9953	9955	9957	9959	9961	9963	9964	9966	9967	9969
2.1	0.9970	9972	9973	9974	9975	9976	9977	9979	9980	9980
2.2	0.9981	9982	9983	9984	9985	9985	9986	9987	9987	9988
2.3	0.9989	9989	9990	9990	9991	9991	9992	9992	9992	9993
2.4	0.9993	9993	9994	9994	9994	9995	9995	9995	9995	9996
2.5	0.9996	9996	9996	9997	9997	9997	9997	9997	9997	9998
2.6	0.9998	9998	9998	9998	9998	9998	9998	9998	9998	9999
2.7	0.9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
2.8	0.9999	9999	9999	9999	9999	9999	9999	0000	0000	0000

x	$\Phi(x)_1$	Diff.	$\Phi(x)_2 : 2$	Diff.	$\Phi(x)_3 : 4$	Diff.
0.00	+ 1.1284		0.0000		− 0.5642	
.01	+ 1.1283	1	− 0.0113	113	− 0.5640	2
.02	1.1279	4	0.0226	113	0.5635	5
.03	1.1274	5	0.0338	112	0.5627	8
.04	+ 1.1266	8	− 0.0451	113	− 0.5615	12
.05	1.1256	10	0.0563	112	0.5600	15
.06	1.1243	13	0.0675	112	0.5581	19
.07	+ 1.1229	14	− 0.0786	111	− 0.5559	22
.08	1.1212	17	0.0897	111	0.5534	25
.09	1.1193	19	0.1007	110	0.5506	28
0.10	+ 1.1172	21	− 0.1117	110	− 0.5474	32
.11	+ 1.1148	24	− 0.1226	109	− 0.5439	35
.12	1.1122	26	0.1335	109	0.5401	38
.13	1.1095	27	0.1442	107	0.5360	41
.14	+ 1.1065	30	− 0.1549	107	− 0.5316	44
.15	1.1033	32	0.1655	106	0.5268	48
.16	1.0999	34	0.1760	105	0.5218	50
.17	+ 1.0962	37	− 0.1864	104	− 0.5164	54
.18	1.0924	38	0.1966	102	0.5108	56
.19	1.0884	40	0.2068	102	0.5049	59
0.20	+ 1.0841	43	− 0.2168	100	− 0.4987	62
.21	+ 1.0797	44	− 0.2267	99	− 0.4922	65
.22	1.0751	46	0.2365	98	0.4855	67
.23	1.0702	49	0.2462	97	0.4785	70
.24	+ 1.0652	50	− 0.2557	95	− 0.4713	72
.25	1.0600	52	0.2650	93	0.4638	75
.26	1.0546	54	0.2742	92	0.4560	78
.27	+ 1.0490	56	− 0.2832	90	− 0.4480	80
.28	1.0433	57	0.2921	89	0.4399	81
.29	1.0374	59	0.3008	87	0.4314	85
0.30	+ 1.0313	61	− 0.3094	86	− 0.4228	86
.31	+ 1.0250	63	− 0.3177	83	− 0.4140	88
.32	1.0186	64	0.3259	82	0.4050	90
.33	1.0120	66	0.3339	80	0.3958	92
.34	+ 1.0052	68	− 0.3418	79	− 0.3864	94
.35	0.9983	69	0.3494	76	0.3769	95
.36	0.9912	71	0.3568	74	0.3671	98
.37	+ 0.9840	72	− 0.3641	73	− 0.3573	98
.38	0.9767	73	0.3711	70	0.3473	100
.39	0.9692	75	0.3780	69	0.3372	101
0.40	+ 0.9615	77	− 0.3846	66	− 0.3269	103
.41	+ 0.9538	77	− 0.3911	65	− 0.3166	103
.42	0.9459	79	0.3973	62	0.3061	105
.43	0.9379	80	0.4033	60	0.2955	106
.44	+ 0.9298	81	− 0.4091	58	− 0.2849	106
.45	0.9215	83	0.4147	56	0.2742	107
.46	0.9132	83	0.4201	54	0.2634	108
.47	+ 0.9047	85	− 0.4252	51	− 0.2525	109
.48	0.8962	85	0.4302	50	0.2416	109
.49	0.8875	87	0.4349	47	0.2307	109
0.50	+ 0.8788	87	− 0.4394	45	− 0.2197	110

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
0.00	0.0000		+ 0.8463		0.0000	
.01	+ 0.0169	169	+ 0.8459	4	− 0.0423	423
.02	0.0338	169	0.8446	13	0.0845	422
.03	0.0507	169	0.8425	21	0.1267	422
.04	+ 0.0675	168	+ 0.8395	30	− 0.1686	419
.05	0.0843	168	0.8357	38	0.2103	417
.06	0.1009	166	0.8311	46	0.2518	415
.07	+ 0.1175	166	+ 0.8257	54	− 0.2928	410
.08	0.1340	165	0.8194	63	0.3335	407
.09	0.1503	163	0.8123	71	0.3737	402
0.10	+ 0.1665	162	+ 0.8045	78	− 0.4134	397
.11	+ 0.1825	160	+ 0.7958	87	− 0.4525	391
.12	0.1983	158	0.7864	94	0.4909	384
.13	0.2139	156	0.7762	102	0.5287	378
.14	+ 0.2293	154	+ 0.7652	110	− 0.5658	371
.15	0.2445	152	0.7535	117	0.6021	363
.16	0.2595	150	0.7411	124	0.6375	354
.17	+ 0.2742	147	+ 0.7280	131	− 0.6721	346
.18	0.2886	144	0.7143	137	0.7057	336
.19	0.3027	141	0.6998	145	0.7384	327
0.20	+ 0.3166	139	+ 0.6847	151	− 0.7701	317
.21	+ 0.3301	135	+ 0.6690	157	− 0.8007	306
.22	0.3433	132	0.6527	163	0.8302	295
.23	0.3562	129	0.6358	169	0.8587	285
.24	+ 0.3688	126	+ 0.6184	174	− 0.8859	272
.25	0.3809	121	0.6004	180	0.9120	261
.26	0.3928	119	0.5819	185	0.9368	248
.27	+ 0.4042	114	+ 0.5629	190	− 0.9604	236
.28	0.4153	111	0.5435	194	0.9827	223
.29	0.4260	107	0.5236	199	1.0038	211
0.30	+ 0.4362	102	+ 0.5034	202	− 1.0235	197
.31	+ 0.4461	99	+ 0.4827	207	− 1.0418	183
.32	0.4555	94	0.4617	210	1.0588	170
.33	0.4646	91	0.4404	213	1.0744	156
.34	+ 0.4731	85	+ 0.4187	217	− 1.0887	143
.35	0.4813	82	0.3968	219	1.1015	128
.36	0.4890	77	0.3747	221	1.1129	114
.37	+ 0.4963	73	+ 0.3523	224	− 1.1229	100
.38	0.5031	68	0.3298	225	1.1315	86
.39	0.5095	64	0.3071	227	1.1387	72
0.40	+ 0.5154	59	+ 0.2842	229	− 1.1445	58
.41	+ 0.5208	54	+ 0.2613	229	− 1.1488	43
.42	0.5258	50	0.2383	230	1.1518	30
.43	0.5304	46	0.2152	231	1.1533	15
.44	+ 0.5344	40	+ 0.1922	230	− 1.1534	1
.45	0.5381	37	0.1691	231	1.1522	12
.46	0.5412	31	0.1461	230	1.1496	26
.47	+ 0.5439	27	+ 0.1231	230	− 1.1457	39
.48	0.5461	22	0.1003	228	1.1404	53
.49	0.5479	18	0.0775	228	1.1338	66
0.50	+ 0.5492	13	+ 0.0549	226	− 1.1259	79

x	$\Phi(x)_1$	Diff.	$\Phi(x)_2 : 2$	Diff.	$\Phi(x)_3 : 4$	Diff.
0.50	+ 0.8788	88	− 0.4394	43	− 0.2197	110
.51	+ 0.8700	90	− 0.4437	40	− 0.2087	110
.52	0.8610	90	0.4477	39	0.1977	110
.53	0.8520	90	0.4516	36	0.1867	110
.54	+ 0.8430	92	− 0.4552	34	− 0.1757	110
.55	0.8338	92	0.4586	32	0.1647	110
.56	0.8246	92	0.4618	30	0.1537	109
.57	+ 0.8154	94	− 0.4648	27	− 0.1428	109
.58	0.8060	93	0.4675	25	0.1319	109
.59	0.7967	95	0.4700	23	0.1210	108
0.60	+ 0.7872	94	− 0.4723	21	− 0.1102	107
.61	+ 0.7778	95	− 0.4744	19	− 0.0995	107
.62	0.7683	96	0.4763	17	0.0888	106
.63	0.7587	96	0.4780	15	0.0782	105
.64	+ 0.7491	96	− 0.4795	12	− 0.0677	104
.65	0.7395	96	0.4807	10	0.0573	103
.66	0.7299	96	0.4817	9	0.0470	102
.67	+ 0.7203	97	− 0.4826	6	− 0.0368	101
.68	0.7106	96	0.4832	5	0.0267	99
.69	0.7010	97	0.4837	2	0.0168	99
0.70	+ 0.6913	97	− 0.4839	0	− 0.0069	97
.71	+ 0.6816	97	− 0.4839	1	+ 0.0028	96
.72	0.6719	97	0.4838	4	0.0124	94
.73	0.6622	96	0.4834	5	0.0218	93
.74	+ 0.6526	97	− 0.4829	7	+ 0.0311	91
.75	0.6429	96	0.4822	9	0.0402	89
.76	0.6333	96	0.4813	11	0.0491	88
.77	+ 0.6237	96	− 0.4802	12	+ 0.0579	87
.78	0.6141	96	0.4790	14	0.0666	84
.79	0.6045	95	0.4776	16	0.0750	83
0.80	+ 0.5950	95	− 0.4760	18	+ 0.0833	81
.81	+ 0.5855	95	− 0.4742	19	+ 0.0914	79
.82	0.5760	94	0.4723	20	0.0993	77
.83	0.5666	94	0.4703	22	0.1070	76
.84	+ 0.5572	93	− 0.4681	24	+ 0.1146	73
.85	0.5479	93	0.4657	25	0.1219	71
.86	0.5386	93	0.4632	27	0.1290	70
.87	+ 0.5293	91	− 0.4605	28	+ 0.1360	67
.88	0.5202	92	0.4577	29	0.1427	66
.89	0.5110	90	0.4548	30	0.1493	63
0.90	+ 0.5020	90	− 0.4518	32	+ 0.1556	61
.91	+ 0.4930	90	− 0.4486	33	+ 0.1617	60
.92	0.4840	88	0.4453	34	0.1677	57
.93	0.4752	88	0.4419	35	0.1734	55
.94	+ 0.4664	88	− 0.4384	37	+ 0.1789	53
.95	0.4576	86	0.4347	37	0.1842	51
.96	0.4490	86	0.4310	38	0.1893	49
.97	+ 0.4404	85	− 0.4272	40	+ 0.1942	46
.98	0.4319	84	0.4232	40	0.1988	45
.99	0.4235	84	0.4192	41	0.2033	43
1.00	+ 0.4151		− 0.4151		+ 0.2076	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
0.50	+ 0.5492		+ 0.0549		− 1.1259	
.51	+ 0.5501	9	+ 0.0325	224	− 1.1168	91
.52	0.5505	4	+ 0.0103	222	1.1064	104
.53	0.5505	0	− 0.0118	221	1.0948	116
.54	+ 0.5501	4	− 0.0335	217	− 1.0820	128
.55	0.5492	9	0.0550	215	1.0681	139
.56	0.5479	13	0.0762	212	1.0530	151
.57	+ 0.5461	18	− 0.0971	209	− 1.0369	161
.58	0.5440	21	0.1177	206	1.0197	172
.59	0.5414	26	0.1379	202	1.0015	182
0.60	+ 0.5385	29	− 0.1578	199	− 0.9823	192
.61	+ 0.5351	34	− 0.1772	194	− 0.9621	202
.62	0.5314	37	0.1962	190	0.9411	210
.63	0.5273	41	0.2148	186	0.9192	219
.64	+ 0.5228	45	− 0.2330	182	− 0.8965	227
.65	0.5180	48	0.2507	177	0.8730	235
.66	0.5128	52	0.2679	172	0.8487	243
.67	+ 0.5072	56	− 0.2846	167	− 0.8238	249
.68	0.5014	58	0.3009	163	0.7982	256
.69	0.4952	62	0.3166	157	0.7720	262
0.70	+ 0.4887	65	− 0.3317	151	− 0.7452	268
.71	+ 0.4819	68	− 0.3464	147	− 0.7180	272
.72	0.4749	70	0.3605	141	0.6902	278
.73	0.4675	74	0.3740	135	0.6621	281
.74	+ 0.4599	76	− 0.3869	129	− 0.6335	286
.75	0.4521	78	0.3993	124	0.6046	289
.76	0.4440	81	0.4111	118	0.5755	291
.77	+ 0.4356	84	− 0.4223	112	− 0.5460	295
.78	0.4271	85	0.4330	107	0.5164	296
.79	0.4183	88	0.4430	100	0.4866	298
0.80	+ 0.4094	89	− 0.4524	94	− 0.4568	298
.81	+ 0.4002	92	− 0.4613	89	− 0.4268	300
.82	0.3909	93	0.4695	82	0.3968	300
.83	0.3814	95	0.4771	76	0.3669	299
.84	+ 0.3718	96	− 0.4842	71	− 0.3369	300
.85	0.3621	97	0.4906	64	0.3071	298
.86	0.3522	99	0.4965	59	0.2774	297
.87	+ 0.3422	100	− 0.5017	52	− 0.2479	295
.88	0.3321	101	0.5064	47	0.2187	292
.89	0.3220	101	0.5105	41	0.1896	291
0.90	+ 0.3117	103	− 0.5140	35	− 0.1609	287
.91	+ 0.3014	103	− 0.5169	29	− 0.1324	285
.92	0.2911	103	0.5193	24	0.1044	280
.93	0.2806	105	0.5211	18	0.0767	277
.94	+ 0.2702	104	− 0.5223	12	− 0.0494	273
.95	0.2598	104	0.5231	8	− 0.0226	268
.96	0.2493	105	0.5232	1	+ 0.0037	263
.97	+ 0.2388	105	− 0.5229	3	+ 0.0296	259
.98	0.2284	104	0.5221	8	0.0549	253
.99	0.2180	104	0.5207	14	0.0796	247
1.00	+ 0.2076	104	− 0.5189	18	+ 0.1038	242

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
1.00	+ 0.4151		— 0.4151		+ 0.2076	
.01	+ 0.4068	83	— 0.4109	42	+ 0.2116	40
.02	0.3987	81	0.4066	43	0.2154	38
.03	0.3906	81	0.4023	43	0.2191	37
.04	+ 0.3826	80	— 0.3979	44	+ 0.2225	34
.05	0.3747	79	0.3934	45	0.2257	32
.06	0.3668	79	0.3889	45	0.2288	31
.07	+ 0.3591	77	— 0.3843	46	+ 0.2316	28
.08	0.3515	76	0.3796	47	0.2342	26
.09	0.3439	76	0.3749	47	0.2367	25
1.10	+ 0.3365	74	— 0.3701	48	+ 0.2389	22
.11	+ 0.3291	74	— 0.3653	48	+ 0.2410	21
.12	0.3219	72	0.3605	48	0.2428	18
.13	0.3147	72	0.3556	49	0.2445	17
.14	+ 0.3076	71	— 0.3507	49	+ 0.2460	15
.15	0.3007	69	0.3458	49	0.2473	13
.16	0.2938	69	0.3408	50	0.2484	11
.17	+ 0.2870	68	— 0.3358	50	+ 0.2494	10
.18	0.2804	66	0.3308	50	0.2502	8
.19	0.2738	66	0.3258	50	0.2508	6
1.20	+ 0.2673	65	— 0.3208	50	+ 0.2513	5
.21	+ 0.2610	63	— 0.3158	50	+ 0.2516	3
.22	0.2547	63	0.3107	51	0.2518	2
.23	0.2485	62	0.3057	50	0.2518	0
.24	+ 0.2425	60	— 0.3007	50	+ 0.2516	2
.25	0.2365	60	0.2957	50	0.2513	3
.26	0.2307	58	0.2906	51	0.2509	4
.27	+ 0.2249	58	— 0.2856	50	+ 0.2503	6
.28	0.2192	57	0.2806	50	0.2496	7
.29	0.2137	55	0.2756	50	0.2487	9
1.30	+ 0.2082	55	— 0.2707	49	+ 0.2478	9
.31	+ 0.2028	54	— 0.2657	50	+ 0.2467	11
.32	0.1976	52	0.2608	49	0.2455	12
.33	0.1924	52	0.2559	49	0.2442	13
.34	+ 0.1873	51	— 0.2510	49	+ 0.2427	15
.35	0.1824	49	0.2462	48	0.2412	15
.36	0.1775	49	0.2414	48	0.2395	17
.37	+ 0.1727	48	— 0.2366	48	+ 0.2378	17
.38	0.1680	47	0.2319	47	0.2360	18
.39	0.1634	46	0.2272	47	0.2341	19
1.40	+ 0.1589	45	— 0.2225	47	+ 0.2321	20
.41	+ 0.1545	44	— 0.2179	46	+ 0.2300	21
.42	0.1502	43	0.2133	46	0.2278	22
.43	0.1460	42	0.2088	45	0.2256	22
.44	+ 0.1419	41	— 0.2043	45	+ 0.2233	23
.45	0.1378	41	0.1999	44	0.2209	24
.46	0.1339	39	0.1955	44	0.2184	25
.47	+ 0.1300	39	— 0.1911	44	+ 0.2159	25
.48	0.1262	38	0.1868	43	0.2134	25
.49	0.1225	37	0.1826	42	0.2108	26
1.50	+ 0.1189	36	— 0.1784	42	+ 0.2081	27

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
1.00	+ 0.2076	104	− 0.5189	23	+ 0.1038	235
.01	+ 0.1972	103	− 0.5166	28	+ 0.1273	230
.02	0.1869	103	0.5138	32	0.1503	223
.03	0.1766	101	0.5106	37	0.1726	216
.04	+ 0.1665	101	− 0.5069	41	+ 0.1942	210
.05	0.1564	100	0.5028	45	0.2152	203
.06	0.1464	100	0.4983	49	0.2355	195
.07	+ 0.1364	98	− 0.4934	53	+ 0.2550	189
.08	0.1266	97	0.4881	57	0.2739	181
.09	0.1169	96	0.4824	60	0.2920	174
1.10	+ 0.1073	94	− 0.4764	63	+ 0.3094	166
.11	+ 0.0979	94	− 0.4701	67	+ 0.3260	159
.12	0.0885	92	0.4634	70	0.3419	151
.13	0.0793	90	0.4564	73	0.3570	144
.14	+ 0.0703	89	− 0.4491	76	+ 0.3714	136
.15	0.0614	88	0.4415	78	0.3850	129
.16	0.0526	86	0.4337	81	0.3979	120
.17	+ 0.0440	84	− 0.4256	83	+ 0.4099	113
.18	0.0356	83	0.4173	85	0.4212	106
.19	0.0273	81	0.4088	87	0.4318	98
1.20	+ 0.0192	79	− 0.4001	90	+ 0.4416	90
.21	+ 0.0113	77	− 0.3911	91	+ 0.4506	83
.22	+ 0.0036	75	0.3820	92	0.4589	75
.23	− 0.0039	74	0.3728	94	0.4664	68
.24	− 0.0113	72	− 0.3634	95	+ 0.4732	61
.25	0.0185	70	0.3539	97	0.4793	53
.26	0.0255	67	0.3442	97	0.4846	47
.27	− 0.0322	66	− 0.3345	98	+ 0.4893	39
.28	0.0388	64	0.3247	99	0.4932	33
.29	0.0452	62	0.3148	100	0.4965	26
1.30	− 0.0514	60	− 0.3048	100	+ 0.4991	19
.31	− 0.0574	58	− 0.2948	100	+ 0.5010	13
.32	0.0632	56	0.2848	101	0.5023	7
.33	0.0688	54	0.2747	101	0.5030	0
.34	− 0.0742	52	− 0.2646	100	+ 0.5030	5
.35	0.0794	50	0.2546	101	0.5025	11
.36	0.0844	48	0.2445	100	0.5014	17
.37	− 0.0892	46	− 0.2345	99	+ 0.4997	23
.38	0.0938	44	0.2246	100	0.4974	27
.39	0.0982	42	0.2146	98	0.4947	33
1.40	− 0.1024	40	− 0.2048	98	+ 0.4914	38
.41	− 0.1064	38	− 0.1950	97	+ 0.4876	42
.42	0.1102	36	0.1853	96	0.4834	47
.43	0.1138	34	0.1757	96	0.4787	51
.44	− 0.1172	32	− 0.1661	94	+ 0.4736	55
.45	0.1204	31	0.1567	93	0.4681	60
.46	0.1235	28	0.1474	92	0.4621	63
.47	− 0.1263	27	− 0.1382	90	+ 0.4558	66
.48	0.1290	25	0.1292	89	0.4492	70
.49	0.1315	23	0.1203	88	0.4422	74
1.50	− 0.1338		− 0.1115		+ 0.4348	

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
1.50	+ 0.1189		− 0.1784		+ 0.2081	
.51	+ 0.1154	35	− 0.1743	41	+ 0.2054	27
.52	0.1120	34	0.1702	41	0.2027	27
.53	0.1086	34	0.1662	40	0.1999	28
.54	+ 0.1053	33	− 0.1622	40	+ 0.1971	28
.55	0.1021	32	0.1583	39	0.1943	28
.56	0.0990	31	0.1544	39	0.1914	29
.57	+ 0.0959	31	− 0.1506	38	+ 0.1885	29
.58	0.0930	29	0.1469	37	0.1856	29
.59	0.0901	29	0.1432	37	0.1826	30
1.60	+ 0.0872	29	− 0.1396	36	+ 0.1797	29
.61	+ 0.0845	27	− 0.1360	36	+ 0.1767	30
.62	0.0818	27	0.1325	35	0.1738	29
.63	0.0792	26	0.1291	34	0.1708	30
.64	+ 0.0766	26	− 0.1257	34	+ 0.1678	30
.65	0.0741	25	0.1223	34	0.1648	30
.66	0.0717	24	0.1191	32	0.1618	30
.67	+ 0.0694	23	− 0.1159	32	+ 0.1588	30
.68	0.0671	23	0.1127	32	0.1558	30
.69	0.0649	22	0.1096	31	0.1528	30
1.70	+ 0.0627	22	− 0.1066	30	+ 0.1499	29
.71	+ 0.0606	21	− 0.1036	30	+ 0.1469	30
.72	0.0586	20	0.1007	29	0.1440	29
.73	0.0566	20	0.0979	28	0.1410	30
.74	+ 0.0546	20	− 0.0951	28	+ 0.1381	29
.75	0.0528	18	0.0924	27	0.1352	29
.76	0.0510	18	0.0897	27	0.1324	28
.77	+ 0.0492	18	− 0.0871	26	+ 0.1295	29
.78	0.0475	17	0.0845	26	0.1267	28
.79	0.0458	17	0.0820	25	0.1239	28
1.80	+ 0.0442	16	− 0.0795	25	+ 0.1211	28
.81	+ 0.0426	16	− 0.0772	23	+ 0.1183	28
.82	0.0411	15	0.0748	24	0.1156	27
.83	0.0396	15	0.0725	23	0.1129	27
.84	+ 0.0382	14	− 0.0703	22	+ 0.1102	27
.85	0.0368	14	0.0681	22	0.1076	26
.86	0.0355	13	0.0660	21	0.1050	26
.87	+ 0.0342	13	− 0.0639	21	+ 0.1024	26
.88	0.0329	13	0.0619	20	0.0999	25
.89	0.0317	12	0.0599	20	0.0974	25
1.90	+ 0.0305	12	− 0.0580	19	+ 0.0949	25
.91	+ 0.0294	11	− 0.0561	19	+ 0.0925	24
.92	0.0283	11	0.0543	18	0.0901	24
.93	0.0272	11	0.0525	18	0.0878	23
.94	+ 0.0262	10	− 0.0508	17	+ 0.0854	24
.95	0.0252	10	0.0491	17	0.0832	22
.96	0.0242	10	0.0475	16	0.0809	23
.97	+ 0.0233	9	− 0.0459	16	+ 0.0787	22
.98	0.0224	9	0.0443	16	0.0765	22
.99	0.0215	9	0.0428	15	0.0744	21
2.00	+ 0.0207	8	− 0.0413	15	+ 0.0723	21

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
1.50	— 0.1338	21	— 0.1115	86	+ 0.4348	76
.51	— 0.1359	20	— 0.1029	85	+ 0.4272	79
.52	0.1379	18	0.0944	83	0.4193	81
.53	0.1397	17	0.0861	81	0.4112	84
.54	— 0.1414	14	— 0.0780	80	+ 0.4028	86
.55	0.1428	14	0.0700	78	0.3942	89
.56	0.1442	11	0.0622	76	0.3853	90
.57	— 0.1453	10	— 0.0546	75	+ 0.3763	91
.58	0.1463	9	0.0471	72	0.3672	93
.59	0.1472	7	0.0399	71	0.3579	95
1.60	— 0.1479	6	— 0.0328	68	+ 0.3484	95
.61	— 0.1485	5	— 0.0260	67	+ 0.3389	97
.62	0.1490	3	0.0193	65	0.3292	97
.63	0.1493	2	0.0128	63	0.3195	99
.64	— 0.1495	1	— 0.0065	61	+ 0.3096	98
.65	0.1496	1	— 0.0004	59	0.2998	99
.66	0.1495	2	+ 0.0055	57	0.2899	99
.67	— 0.1493	2	+ 0.0112	55	+ 0.2800	99
.68	0.1491	4	0.0167	53	0.2701	99
.69	0.1487	5	0.0220	51	0.2602	99
1.70	— 0.1482	6	+ 0.0271	49	+ 0.2503	98
.71	— 0.1476	7	+ 0.0320	47	+ 0.2405	98
.72	0.1469	8	0.0367	45	0.2307	98
.73	0.1461	8	0.0412	44	0.2209	96
.74	— 0.1453	10	+ 0.0456	41	+ 0.2113	96
.75	0.1443	10	0.0497	39	0.2017	95
.76	0.1433	11	0.0536	38	0.1922	94
.77	— 0.1422	12	+ 0.0574	35	+ 0.1828	93
.78	0.1410	13	0.0609	34	0.1735	92
.79	0.1397	13	0.0643	32	0.1643	90
1.80	— 0.1384	14	+ 0.0675	30	+ 0.1553	89
.81	— 0.1370	14	+ 0.0705	29	+ 0.1464	87
.82	0.1356	15	0.0734	26	0.1377	86
.83	0.1341	16	0.0760	25	0.1291	85
.84	— 0.1325	15	+ 0.0785	24	+ 0.1206	83
.85	0.1310	17	0.0809	21	0.1123	81
.86	0.1293	17	0.0830	20	0.1042	79
.87	— 0.1276	17	+ 0.0850	19	+ 0.0963	78
.88	0.1259	17	0.0869	17	0.0885	76
.89	0.1242	18	0.0886	15	0.0809	74
1.90	— 0.1224	18	+ 0.0901	14	+ 0.0735	72
.91	— 0.1206	19	+ 0.0915	13	+ 0.0663	70
.92	0.1187	19	0.0928	11	0.0593	68
.93	0.1168	18	0.0939	10	0.0525	66
.94	— 0.1150	19	+ 0.0949	8	+ 0.0459	64
.95	0.1131	20	0.0957	7	0.0395	63
.96	0.1111	19	0.0964	7	0.0332	60
.97	— 0.1092	19	+ 0.0971	4	+ 0.0272	58
.98	0.1073	20	0.0975	4	0.0214	56
.99	0.1053	20	0.0979	3	0.0158	55
2.00	— 0.1033		+ 0.0982		+ 0.0103	

x	$\Phi(x)_1$	$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
2.00	+ 0.0207	− 0.0413	14	+ 0.0723	20
.01	+ 0.0199	− 0.0399	14	+ 0.0703	20
.02	0.0191	0.0385	13	0.0683	20
.03	0.0183	0.0372	13	0.0663	19
.04	+ 0.0176	− 0.0359	13	+ 0.0644	19
.05	0.0169	0.0346	12	0.0625	19
.06	0.0162	0.0334	12	0.0606	18
.07	+ 0.0155	− 0.0322	12	+ 0.0588	17
.08	0.0149	0.0310	11	0.0571	18
.09	0.0143	0.0299	11	0.0553	17
2.10	+ 0.0137	− 0.0288	11	+ 0.0536	16
.11	+ 0.0132	− 0.0277	10	+ 0.0520	16
.12	0.0126	0.0267	10	0.0504	16
.13	0.0121	0.0257	9	0.0488	16
.14	+ 0.0116	− 0.0248	10	+ 0.0472	15
.15	0.0111	0.0238	9	0.0457	15
.16	0.0106	0.0229	8	0.0442	14
.17	+ 0.0102	− 0.0221	9	+ 0.0428	14
.18	0.0097	0.0212	8	0.0414	13
.19	0.0093	0.0204	8	0.0401	14
2.20	+ 0.0089	− 0.0196	7	+ 0.0387	13
.21	+ 0.0085	− 0.0189	8	+ 0.0374	12
.22	0.0082	0.0181	7	0.0362	13
.23	0.0078	0.0174	7	0.0349	12
.24	+ 0.0075	− 0.0167	6	+ 0.0337	11
.25	0.0071	0.0161	7	0.0326	11
.26	0.0068	0.0154	6	0.0315	11
.27	+ 0.0065	− 0.0148	6	+ 0.0304	11
.28	0.0062	0.0142	6	0.0293	10
.29	0.0060	0.0136	5	0.0283	10
2.30	+ 0.0057	− 0.0131	6	+ 0.0273	10
.31	+ 0.0054	− 0.0125	5	+ 0.0263	10
.32	0.0052	0.0120	5	0.0253	9
.33	0.0050	0.0115	4	0.0244	9
.34	+ 0.0047	− 0.0111	5	+ 0.0235	9
.35	0.0045	0.0106	4	0.0226	8
.36	0.0043	0.0102	5	0.0218	8
.37	+ 0.0041	− 0.0097	4	+ 0.0210	8
.38	0.0039	0.0093	4	0.0202	8
.39	0.0037	0.0089	4	0.0194	7
2.40	+ 0.0036	− 0.0085	3	+ 0.0187	7
.41	+ 0.0034	− 0.0082	4	+ 0.0180	7
.42	0.0032	0.0078	3	0.0173	7
.43	0.0031	0.0075	4	0.0166	6
.44	+ 0.0029	− 0.0071	3	+ 0.0160	6
.45	0.0028	0.0068	3	0.0154	7
.46	0.0027	0.0065	3	0.0147	5
.47	+ 0.0025	− 0.0062	2	+ 0.0142	6
.48	0.0024	0.0060	3	0.0136	5
.49	0.0023	0.0057	3	0.0131	6
2.50	+ 0.0022	− 0.0054		+ 0.0125	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
2.00	— 0.1033	19	+ 0.0982	1	+ 0.0103	52
.01	— 0.1014	20	+ 0.0983	1	+ 0.0051	50
.02	0.0994	20	0.0984	1	+ 0.0001	48
.03	0.0974	19	0.0983	1	— 0.0047	47
.04	— 0.0955	20	+ 0.0982	2	— 0.0094	44
.05	0.0935	19	0.0980	4	0.0138	42
.06	0.0916	20	0.0976	4	0.0180	41
.07	— 0.0896	19	+ 0.0972	4	— 0.0221	38
.08	0.0877	20	0.0968	6	0.0259	37
.09	0.0857	19	0.0962	6	0.0296	35
2.10	— 0.0838	19	+ 0.0956	7	— 0.0331	33
.11	— 0.0819	19	+ 0.0949	8	— 0.0364	32
.12	0.0800	19	0.0941	8	0.0396	28
.13	0.0781	18	0.0933	9	0.0424	28
.14	— 0.0763	18	+ 0.0924	9	— 0.0452	26
.15	0.0745	19	0.0915	10	0.0478	24
.16	0.0726	18	0.0905	10	0.0502	23
.17	— 0.0708	17	+ 0.0895	11	— 0.0525	21
.18	0.0691	18	0.0884	11	0.0546	20
.19	0.0673	17	0.0873	12	0.0566	18
2.20	— 0.0656	18	+ 0.0861	11	— 0.0584	17
.21	— 0.0638	16	+ 0.0850	13	— 0.0601	15
.22	0.0622	17	0.0837	12	0.0616	14
.23	0.0605	16	0.0825	13	0.0630	12
.24	— 0.0589	16	+ 0.0812	13	— 0.0642	11
.25	0.0573	16	0.0799	13	0.0653	10
.26	0.0557	16	0.0786	13	0.0663	9
.27	— 0.0541	15	+ 0.0773	14	— 0.0672	8
.28	0.0526	15	0.0759	13	0.0680	6
.29	0.0511	15	0.0746	14	0.0686	5
2.30	— 0.0496	15	+ 0.0732	14	— 0.0691	5
.31	— 0.0481	14	+ 0.0718	14	— 0.0696	3
.32	0.0467	14	0.0704	14	0.0699	2
.33	0.0453	13	0.0690	14	0.0701	2
.34	— 0.0440	14	+ 0.0676	14	— 0.0703	0
.35	0.0426	13	0.0662	14	0.0703	0
.36	0.0413	13	0.0648	14	0.0703	1
.37	— 0.0400	12	+ 0.0634	14	— 0.0702	2
.38	0.0388	12	0.0620	14	0.0700	3
.39	0.0376	12	0.0606	14	0.0697	3
2.40	— 0.0364	12	+ 0.0592	14	— 0.0694	4
.41	— 0.0352	12	+ 0.0578	14	— 0.0690	5
.42	0.0340	11	0.0564	13	0.0685	5
.43	0.0329	11	0.0551	14	0.0680	6
.44	— 0.0318	10	+ 0.0537	13	— 0.0674	6
.45	0.0308	11	0.0524	14	0.0668	7
.46	0.0297	10	0.0510	13	0.0661	7
.47	— 0.0287	9	+ 0.0497	13	— 0.0654	8
.48	0.0278	10	0.0484	13	0.0646	8
.49	0.0268	9	0.0471	12	0.0638	8
2.50	— 0.0259		+ 0.0459		— 0.0630	

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$		$\Phi(x)_5:16$		$\Phi(x)_6:32$	
2.50	+ 0.0022	− 0.0054	+ 0.0125	− 0.0259	9	+ 0.0459	13	− 0.0630	9
.51	+ 0.0021	− 0.0052	+ 0.0120	− 0.0250	9	+ 0.0446	12	− 0.0621	9
.52	0.0020	0.0050	0.0115	0.0241	9	0.0434	12	0.0612	9
.53	0.0019	0.0047	0.0111	0.0232	8	0.0422	12	0.0603	10
.54	+ 0.0018	− 0.0045	+ 0.0106	− 0.0224	8	+ 0.0410	12	− 0.0593	10
.55	0.0017	0.0043	0.0102	0.0216	8	0.0398	11	0.0583	10
.56	0.0016	0.0041	0.0097	0.0208	8	0.0387	12	0.0573	10
.57	+ 0.0015	− 0.0039	+ 0.0093	− 0.0200	7	+ 0.0375	11	− 0.0563	10
.58	0.0015	0.0037	0.0089	0.0193	7	0.0364	11	0.0553	10
.59	0.0014	0.0036	0.0086	0.0186	7	0.0353	11	0.0543	11
2.60	+ 0.0013	− 0.0034	+ 0.0082	− 0.0179	7	+ 0.0342	10	− 0.0532	10
.61	+ 0.0012	− 0.0032	+ 0.0078	− 0.0172	6	+ 0.0332	11	− 0.0522	11
.62	0.0012	0.0031	0.0075	0.0166	7	0.0321	10	0.0511	11
.63	0.0011	0.0029	0.0072	0.0159	6	0.0311	10	0.0500	11
.64	+ 0.0011	− 0.0028	+ 0.0069	− 0.0153	6	+ 0.0301	9	− 0.0489	10
.65	0.0010	0.0027	0.0066	0.0147	6	0.0292	10	0.0479	11
.66	0.0010	0.0025	0.0063	0.0141	5	0.0282	9	0.0468	11
.67	+ 0.0009	− 0.0024	+ 0.0060	− 0.0136	5	+ 0.0273	9	− 0.0457	11
.68	0.0009	0.0023	0.0057	0.0131	6	0.0264	9	0.0446	10
.69	0.0008	0.0022	0.0055	0.0125	5	0.0255	8	0.0436	11
2.70	+ 0.0008	− 0.0021	+ 0.0052	− 0.0120	4	+ 0.0247	9	− 0.0425	11
.71	+ 0.0007	− 0.0020	+ 0.0050	− 0.0116	5	+ 0.0238	8	− 0.0414	10
.72	0.0007	0.0019	0.0048	0.0111	5	0.0230	8	0.0404	11
.73	0.0007	0.0018	0.0045	0.0106	4	0.0222	8	0.0393	10
.74	+ 0.0006	− 0.0017	+ 0.0043	− 0.0102	4	+ 0.0214	7	− 0.0383	10
.75	0.0006	0.0016	0.0041	0.0098	4	0.0207	8	0.0373	10
.76	0.0006	0.0015	0.0039	0.0094	4	0.0199	7	0.0363	10
.77	+ 0.0005	− 0.0015	+ 0.0038	− 0.0090	4	+ 0.0192	7	− 0.0353	10
.78	0.0005	0.0014	0.0036	0.0086	4	0.0185	7	0.0343	10
.79	0.0005	0.0013	0.0034	0.0082	3	0.0178	6	0.0333	9
2.80	+ 0.0004	− 0.0012	+ 0.0033	− 0.0079	4	+ 0.0172	6	− 0.0324	10
.81	+ 0.0004	− 0.0012	+ 0.0031	− 0.0075	3	+ 0.0166	7	− 0.0314	9
.82	0.0004	0.0011	0.0030	0.0072	3	0.0159	6	0.0305	9
.83	0.0004	0.0011	0.0028	0.0069	3	0.0153	6	0.0296	9
.84	+ 0.0004	− 0.0010	+ 0.0027	− 0.0066	3	+ 0.0147	5	− 0.0287	9
.85	0.0003	0.0010	0.0026	0.0063	3	0.0142	6	0.0278	9
.86	0.0003	0.0009	0.0024	0.0060	2	0.0136	5	0.0269	8
.87	+ 0.0003	− 0.0009	+ 0.0023	− 0.0058	3	+ 0.0131	5	− 0.0261	9
.88	0.0003	0.0008	0.0022	0.0055	2	0.0126	5	0.0252	8
.89	0.0003	0.0008	0.0021	0.0053	3	0.0121	5	0.0244	8
2.90	+ 0.0003	− 0.0007	+ 0.0020	− 0.0050	2	+ 0.0116	4	− 0.0236	8
.91	+ 0.0002	− 0.0007	+ 0.0019	− 0.0048	2	+ 0.0112	5	− 0.0228	7
.92	0.0002	0.0007	0.0018	0.0046	2	0.0107	4	0.0221	8
.93	0.0002	0.0006	0.0017	0.0044	2	0.0103	4	0.0213	7
.94	+ 0.0002	− 0.0006	+ 0.0016	− 0.0042	2	+ 0.0099	5	− 0.0206	7
.95	0.0002	0.0006	0.0015	0.0040	2	0.0094	3	0.0199	7
.96	0.0002	0.0005	0.0015	0.0038	2	0.0091	4	0.0192	7
.97	+ 0.0002	− 0.0005	+ 0.0014	− 0.0036	1	+ 0.0087	4	− 0.0185	6
.98	0.0002	0.0005	0.0013	0.0035	2	0.0083	3	0.0179	7
.99	0.0001	0.0004	0.0012	0.0033	2	0.0080	4	0.0172	6
3.00	+ 0.0001	− 0.0004	+ 0.0012	− 0.0031		+ 0.0076		− 0.0166	

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$	$\Phi(x)_5:16$	$\Phi(x)_6:32$
3.00	+ 0.0001	− 0.0004	+ 0.0012	− 0.0031	+ 0.0076	− 0.0166
.01	+ 0.0001	− 0.0004	+ 0.0011	− 0.0030	+ 0.0073	− 0.0160
.02	0.0001	0.0004	0.0011	0.0028	0.0070	0.0154
.03	0.0001	0.0004	0.0010	0.0027	0.0067	0.0148
.04	+ 0.0001	− 0.0003	+ 0.0010	− 0.0026	+ 0.0064	− 0.0143
.05	0.0001	0.0003	0.0009	0.0024	0.0061	0.0137
.06	0.0001	0.0003	0.0009	0.0023	0.0058	0.0132
.07	+ 0.0001	− 0.0003	+ 0.0008	− 0.0022	+ 0.0056	− 0.0127
.08	0.0001	0.0003	0.0008	0.0021	0.0053	0.0122
.09	0.0001	0.0002	0.0007	0.0020	0.0051	0.0117
3.10	+ 0.0001	− 0.0002	+ 0.0007	− 0.0019	+ 0.0049	− 0.0113
.11	+ 0.0001	− 0.0002	+ 0.0007	− 0.0018	+ 0.0046	− 0.0108
.12	0.0001	0.0002	0.0006	0.0017	0.0044	0.0104
.13	0.0001	0.0002	0.0006	0.0016	0.0042	0.0100
.14	+ 0.0001	− 0.0002	+ 0.0006	− 0.0015	+ 0.0040	− 0.0096
.15	0.0001	0.0002	0.0005	0.0015	0.0038	0.0092
.16	0.0001	0.0002	0.0005	0.0014	0.0037	0.0088
.17	+ 0.0000	− 0.0002	+ 0.0005	− 0.0013	+ 0.0035	− 0.0084
.18		0.0001	0.0004	0.0013	0.0033	0.0081
.19		0.0001	0.0004	0.0012	0.0032	0.0077
3.20		− 0.0001	+ 0.0004	− 0.0011	+ 0.0030	− 0.0074
.21		− 0.0001	+ 0.0004	− 0.0011	+ 0.0029	− 0.0071
.22		0.0001	0.0003	0.0010	0.0027	0.0068
.23		0.0001	0.0003	0.0010	0.0026	0.0065
.24		− 0.0001	+ 0.0003	− 0.0009	+ 0.0025	− 0.0062
.25		0.0001	0.0003	0.0009	0.0024	0.0059
.26		0.0001	0.0003	0.0008	0.0022	0.0057
.27		− 0.0001	+ 0.0003	− 0.0008	+ 0.0021	− 0.0054
.28		0.0001	0.0002	0.0007	0.0020	0.0052
.29		0.0001	0.0002	0.0007	0.0019	0.0049
3.30		− 0.0001	+ 0.0002	− 0.0007	+ 0.0018	− 0.0047
.31		− 0.0001	+ 0.0002	− 0.0006	+ 0.0017	− 0.0045
.32		0.0001	0.0002	0.0006	0.0016	0.0043
.33		0.0001	0.0002	0.0006	0.0016	0.0041
.34		− 0.0001	+ 0.0002	− 0.0005	+ 0.0015	− 0.0039
.35		0.0001	0.0002	0.0005	0.0014	0.0037
.36		0.0000	0.0002	0.0005	0.0013	0.0035
.37			+ 0.0001	− 0.0004	+ 0.0013	− 0.0034
.38			0.0001	0.0004	0.0012	0.0032
.39			0.0001	0.0004	0.0011	0.0031
3.40			+ 0.0001	− 0.0004	+ 0.0011	− 0.0029
.41			+ 0.0001	− 0.0003	+ 0.0010	− 0.0028
.42			0.0001	0.0003	0.0010	0.0026
.43			0.0001	0.0003	0.0009	0.0025
.44			+ 0.0001	− 0.0003	+ 0.0009	− 0.0024
.45			0.0001	0.0003	0.0008	0.0023
.46			0.0001	0.0003	0.0008	0.0022
.47			+ 0.0001	− 0.0002	+ 0.0007	− 0.0020
.48			0.0001	0.0002	0.0007	0.0019
.49			0.0001	0.0002	0.0007	0.0018
3.50			+ 0.0001	− 0.0002	+ 0.0006	− 0.0017

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$	$\Phi(x)_5:16$	$\Phi(x)_6:32$
3.50			+ 0.0001	− 0.0002	+ 0.0006	− 0.0017
.51			+ 0.0001	− 0.0002	+ 0.0006	− 0.0017
.52			0.0001	0.0002	0.0005	0.0016
.53			0.0001	0.0002	0.0005	0.0015
.54			0.0000	− 0.0002	+ 0.0005	− 0.0014
.55				0.0001	0.0005	0.0013
.56				0.0001	0.0004	0.0013
.57				− 0.0001	+ 0.0004	− 0.0012
.58				0.0001	0.0004	0.0011
.59				0.0001	0.0004	0.0011
3.60				− 0.0001	+ 0.0003	− 0.0010
.61				− 0.0001	+ 0.0003	− 0.0010
.62				0.0001	0.0003	0.0009
.63				0.0001	0.0003	0.0009
.64				− 0.0001	+ 0.0003	− 0.0008
.65				0.0001	0.0003	0.0008
.66				0.0001	0.0002	0.0007
.67				− 0.0001	+ 0.0002	− 0.0007
.68				0.0001	0.0002	0.0007
.69				0.0001	0.0002	0.0006
3.70				− 0.0001	+ 0.0002	− 0.0006
.71				− 0.0001	+ 0.0002	− 0.0005
.72				0.0001	0.0002	0.0005
.73				0.0000	0.0002	0.0005
.74					+ 0.0001	− 0.0005
.75					0.0001	0.0004
.76					0.0001	0.0004
.77					+ 0.0001	− 0.0004
.78					0.0001	0.0004
.79					0.0001	0.0003
3.80					+ 0.0001	− 0.0003
.81					+ 0.0001	− 0.0003
.82					0.0001	0.0003
.83					0.0001	0.0003
.84					+ 0.0001	− 0.0003
.85					0.0001	0.0002
.86					0.0001	0.0002
.87					+ 0.0001	− 0.0002
.88					0.0001	0.0002
.89					0.0001	0.0002
3.90					+ 0.0001	− 0.0002
.91					0.0000	− 0.0002
.92						0.0002
.93						0.0001
.94						− 0.0001
.95						0.0001
.96						0.0001
.97						− 0.0001
.98						0.0001
.99						0.0001
4.00						− 0.0001

Sachregister.

(Die Zahlen beziehen sich auf die Seiten.)

Namenregister.

(Die Zahlen beziehen sich auf die Seiten.)

Blaschke, Dr. **E.**, Regierungsrat im Minist. d. Innern, Professor an der Technischen Hochschule zu Wien, Vorlesungen über mathematische Statistik. Die Lehre von den statistischen Maßzahlen. Mit 17 Textfiguren und 5 Tafeln. [VIII u. 268 S.] gr. 8. 1906. In Leinwand geb. n. ℳ 7.40.

Im ersten Teile der Vorlesungen — die zunächst als Studienbehelf für die Hörer der an den Hochschulen bestehenden Kurse für Versicherungstechnik dienen, dann aber auch angesichts der außerordentlichen Entwicklung des Personenversicherungswesens den Interessen weiterer Kreise entgegenkommen sollen — werden die Methoden zur Herstellung einwandfreier statistischen Tabellen (Absterbeordnungen, Invaliditätstafeln, Krankentafeln, Heiratsordnungen usw.), im zweiten Teile auf Grundlage von Untersuchungen über die Bedeutung der Tabellen die Anwendungen erörtert, welche sich hieraus einerseits für die Theorie der Personenversicherung, andererseits für das unter dem Namen der Tafelausgleichung bekannte statistische Problem ergeben.

Das Bestreben, den Wissenszweig vor allem für die Praxis nutzbar zu machen, führte dazu, wenn auch nur im Anhange, auf einige mechanische Hilfsmittel der Forschung (die Zählmaschine) hinzuweisen und damit für deren allgemeine Verwendung die Wege zu ebnen.

Bortkewitsch, Dr. L. von, Professor an der Universität Berlin, das Gesetz der kleinen Zahlen. [VII u. 52 S.] gr. 8. 1898. geh. n. ℳ 2.—

Die vorliegende Abhandlung stellt den ersten Versuch dar, statistischen Reihen, welche aus kleinen absoluten Zahlen bestehen, vom Standpunkte der Wahrscheinlichkeitsrechnung aus näher zu treten, d. h. die Gesetze des Zufalls an diesen statistischen Daten zu untersuchen. Es ergibt sich, daß die bei den untersuchten Reihen gefundenen Schwankungen den Voraussagungen der Theorie fast vollständig entsprechen, worin eben das Gesetz der kleinen Zahlen besteht. Den scheinbaren Widerspruch zwischen dem Verhalten der großen und kleinen Ereigniszahlen erklärt Lexis Theorie der Fehlerexzedenten.

Bruns, Geheimer Hofrat Dr. **Heinrich**, Professor an der Universität Leipzig, Grundlinien des wissenschaftlichen Rechnens. [VI u. 159 S.] gr. 8. 1903. geh. n. ℳ 3.40, in Leinwand geb. n. ℳ 4.—

Der Verfasser hatte bei den Übungen in seinem Seminar für „wissenschaftliches Rechnen" schon vor längerer Zeit damit begonnen, den Teilnehmern die zur Vorbereitung erforderlichen mathematischen Entwicklungen autographiert in die Hand zu geben, um dadurch Zeit für die Behandlung besonderer Aufgaben zu gewinnen. Diese Aufzeichnungen werden hier in etwas erweiterter Gestalt der Öffentlichkeit übergeben, da es sich um Dinge handelt, für die es bisher an einer handlichen Zusammenstellung fehlte, und die überdies außerhalb des Kreises der berufsmäßigen Rechner keineswegs so bekannt sind, wie sie es bei ihrer erprobten Nützlichkeit verdienen.

——— Wahrscheinlichkeitsrechnung und Kollektivmaßlehre. [VIII u. 310 S. und Anhang 18 S.] gr. 8. 1906. geh. n. ℳ 7.80, in Leinwand geb. n. ℳ 8.40.

Das vorliegende Buch ist aus den Vorlesungen entstanden, die der Verfasser über Wahrscheinlichkeitsrechnung gehalten hat. Nachdem es ihm gelungen war, eine brauchbare analytische Darstellung für willkürliche Verteilungsfunktionen aufzufinden, hat er das Hauptgewicht auf die Entwicklung der Kollektivmaßlehre gelegt, die rund zwei Drittel des Werkes ausfüllt. Die sogenannten Anwendungen der Wahrscheinlichkeitsrechnung auf Versicherungswesen, Statistik und Fehlertheorie werden nur flüchtig gestreift, weil diese Gegenstände längst über den Rahmen einer bloßen Anwendung hinausgewachsen sind und eine selbständige Behandlung beanspruchen dürfen. Als Ausgleich bietet das Werk die erste lehrbuchmäßige Darstellung der allgemeinen Kollektivmaßlehre.

Cantor, Geheimer Hofrat Dr. **M.**, Professor an der Universität Heidelberg, politische Arithmetik oder die Arithmetik des täglichen Lebens. 2. Auflage. [X u. 155 S.] gr. 8. 1903. In Leinwand geb. n. ℳ 1.80.

Das vorliegende Büchlein umfaßt den Inhalt der Vorlesungen, die der Verfasser allwinterlich über politische Arithmetik für Kameralisten an der Universität Heidelberg hält und zu deren Veröffentlichung er sich entschloß, weil keine der vorhandenen Schriften entfernt den gleichen Stoff, der hier auf ca. 10 Bogen zusammengedrängt ist, behandelt.






Stanford University Libraries
3 6105 002 056 229